

Shrinkage estimation of a mean matrix of a multivariate complex normal distribution

Yoshihiko Konno *

February 11, 2013

Abstract

The problem of estimating a mean matrix of a multivariate complex normal distribution with an unknown covariance matrix is considered under an invariant loss function. By using complex versions of the Stein identity, the Stein-Haff identity, and calculus on eigenvalues, a formula is obtained for an unbiased estimate of the risk of an invariant class of estimators, from which several minimax shrinkage estimators are constructed.

1 Introduction

The multivariate complex normal and complex Wishart distributions were first explored in Goodman [14], and followed by Khatri [20]. These models play an important role in signal processing methods. See Kay [19] for the need of complex data models and DoGondžić and Neborai [11] for a unified approach based on complex GMANOVA models to analyze and extend signal processing models. See Ratnarajah *et al.* [32], Micheas *et al.* [30], and Díaz-García and Gutierrez-Jáimez [10] for recent development of complex data model. Lillestøl [26] first investigated Stein-like shrinkage methods on simultaneous estimation of a mean vector

2000 *Mathematics Subject Classification.* Primary: 62H12, Secondary: 62F10.

Key words and phrases. unbiased estimation of risk, integration-by-parts identities, complex Wishart distribution

of the complex normal model. However, shrinkage methods for these models have received less attention so far, although it is important to develop these methods beyond the maximum likelihood estimator of estimating the unknown signals in the multivariate complex normal distribution. The goal of this paper is to show how certain decision theoretical results concerning the problem of estimating a mean matrix of the real normal distribution can be extended to the complex multivariate normal case.

In this paper, we consider the problem of estimating an $m \times p$ unknown constant complex matrix $\mathbf{\Xi}$ that is observed with additive complex normal random errors in a decision theoretic set-up. Our observations are an $m \times p$ data matrix \mathbf{Z} and a $p \times p$ positive definite Hermitian matrix \mathbf{S} , which is represented as

$$\begin{aligned} \mathbf{Z} : m \times p &\sim \mathbb{CN}_{m \times p}(\mathbf{\Xi}, \mathbf{K} \otimes \mathbf{\Sigma}), \\ \mathbf{S} : p \times p &\sim \mathbb{CW}_p(\mathbf{\Sigma}, n) \quad \text{with } \mathbf{Z} \text{ and } \mathbf{S} \text{ independent,} \end{aligned} \tag{1}$$

where $n > p$, $\mathbf{\Sigma}$ is a $p \times p$ positive definite Hermitian constant matrix, and \mathbf{K} is an $m \times m$ positive definite Hermitian constant matrix. Here we assume that $\mathbf{\Xi}$ and $\mathbf{\Sigma}$ are unknown although we assume that \mathbf{K} is known. Furthermore $\mathbb{CN}_{m \times p}(\mathbf{\Xi}, \mathbf{K} \otimes \mathbf{\Sigma})$ and $\mathbb{CW}_p(\mathbf{\Sigma}, n)$ stand for a matrix-variate complex normal distribution with the mean matrix $\mathbf{\Xi}$ and the covariance matrix $\mathbf{K} \otimes \mathbf{\Sigma}$ and a complex Wishart distribution with the degree of freedom n and the parameters $\mathbf{\Sigma}$, respectively. In other words, the model (1) means that the density of $\mathbf{K}^{-1/2}\mathbf{Z} =: \tilde{\mathbf{Z}}$ with respect to the Lebesgue measure on $\mathbb{C}^{m \times p}$ is given as

$$\pi^{-mp} \text{Det}(\mathbf{\Sigma})^{-m} \exp\{-\text{Tr}((\mathbf{z} - \tilde{\mathbf{\Xi}})\mathbf{\Sigma}^{-1}(\mathbf{z} - \tilde{\mathbf{\Xi}})^*)\}, \quad \mathbf{z} \in \mathbb{C}^{m \times p},$$

where $\tilde{\mathbf{\Xi}} = \mathbf{K}^{-1/2}\mathbf{\Xi}$, while the density of \mathbf{S} with respect to the Lebesgue measure on $\mathbb{C}_+^{p \times p}$ is given by

$$\frac{\text{Det}(\mathbf{s})^{n-p} \exp(-\text{Tr}(\mathbf{s}\mathbf{\Sigma}^{-1}))}{\text{Det}(\mathbf{\Sigma})^n \pi^{p(p-1)/2} \prod_{k=1}^p \Gamma(n+1-k)}, \quad \mathbf{s} \in \mathbb{C}_+^{p \times p}. \tag{2}$$

Here $\Gamma(\cdot)$ is the usual Gamma function, $\text{Tr}(\cdot)$ and $\text{Det}(\cdot)$ denote the trace and determinant of a square matrix, and the superscript “*” means the complex conjugate transpose of a matrix. Furthermore $\mathbb{C}^{m \times p}$ and $\mathbb{C}_+^{p \times p}$ stand for the sets of all $m \times p$ complex matrices and of all $p \times p$ positive definite Hermitian complex matrices, respectively.

Based on (\mathbf{Z}, \mathbf{S}) we consider the problem of estimating the mean matrix $\mathbf{\Xi}$ with respect to a loss function

$$\mathcal{L}(\hat{\mathbf{\Xi}}, (\mathbf{\Xi}, \mathbf{\Sigma})) = \text{Tr} \{ \mathbf{\Sigma}^{-1} (\hat{\mathbf{\Xi}} - \mathbf{\Xi})^* \mathbf{K}^{-1} (\hat{\mathbf{\Xi}} - \mathbf{\Xi}) \},$$

where an $m \times p$ random matrix $\hat{\mathbf{\Xi}}$ is an estimator of $\mathbf{\Xi}$. The risk function corresponding to this loss function is

$$\mathcal{R}(\hat{\mathbf{\Xi}}, (\mathbf{\Xi}, \mathbf{\Sigma})) = \mathbb{E}[\mathcal{L}(\hat{\mathbf{\Xi}}, (\mathbf{\Xi}, \mathbf{\Sigma}))],$$

where the expectation above is taken with respect to the joint distribution of (\mathbf{Z}, \mathbf{S}) .

This estimation problem is important since it is a prototype of estimating the regression matrix of a complex MANOVA model and of predicting multivariate responses in a linear regression complex model. We extend a large body of the results obtained by Efron and Morris [12], Bilodeau and Kariya [5], Kariya *et al.* [18], Konno [21], and van der Merwe and Zidek [38] in the multivariate real normal set-up to the complex normal set-up (1). The results in the real normal model were obtained by extensive use of the integration by parts approach, known as the Stein identity derived by Stein [34, 36], and the Stein-Haff identity by Stein [35] and Haff [15, 16]. In addition to these identities, the eigenvalue calculus, developed by Loh [27, 28, 29], Konno [21], and Kariya *et al.* [18], is important to the development for a systematic search for shrinkage estimators. We extend these approaches to the complex normal set-up. The Stein identity for the multivariate complex normal is easily derived by using an isomorphism between real and complex variables stated in Andersen *et al.* [1] while the Stein-Haff identity was extended to the complex Wishart distribution by Svensson and Lundberg [37]. These identities and the eigenvalue calculus for the complex matrix developed in this paper are exploited to establish a systematic search for shrinkage estimators for the model (1), which includes the FICYREG estimator of van der Merwe and Zidek [38].

Shrinkage methods for estimating the regression matrix in a multivariate linear regression model have been extensively investigated to overcome the shortcomings of the ordinary least squares estimator. The literature includes Brown and Zidek [7, 8], Dempster [9], and van der Merwe and Zidek [38]. Later Breiman and Friedman [6] proposed to predict a future observation by a ridge-type shrinkage estimator in order to use information of correlated variables. See also

Bilodeau [4], Oman [31], and Srivastava and Solanky [33] for further investigation on this problem. As mentioned in Srivastava and Solanky [33], we can use minimax estimators to construct better predictors in order to overcome shortcomings of the predictor based on the least squares estimator. This shows that the results obtained in this paper can be immediately applied to the problem of predicting a future observation in a multivariate linear model for complex data.

The remaining parts of this papers are organized as follows. In Section 2, we state some notation and the integration by parts formulae. In Section 3, we develop shrinkage estimators for the known covariance case, which is an extension to the results obtained in Stein [35] and Zheng [39, 40]. In Section 4, we obtain unbiased risk estimate for invariant estimators, from which several shrinkage estimators are derived. In the Appendix, the results on eigenvalue calculus for the complex set-up and their proofs are developed.

2 Preliminaries: Notation and Basic identities

This section first presents some notation used throughout this paper. Next we introduce integration by parts formulae, complex versions of the Stein identity and the Stein-Haff identity, which play vital roles in obtaining unbiased risk estimate in Sections 3 and 4.

2.1 Notation

Let \mathbb{R} and \mathbb{C} denote the field of real and complex numbers, respectively. We represent any element $c \in \mathbb{C}$ as $c = a + \sqrt{-1}b$, where $a, b \in \mathbb{R}$. We also denote the real and imaginary parts of c by $\text{Re } c$ and $\text{Im } c$, respectively. In particular we denote by \mathbb{R}_+ the set of all positive real numbers. The conjugate of a complex number c is given by $\bar{c} := a - \sqrt{-1}b$. We define by \mathbb{R}^p and \mathbb{C}^p the sets of all p -tuples of real and complex numbers, respectively. We set $\mathbb{R}_{>}^p = \{(\ell_1, \ell_2, \dots, \ell_p) \in \mathbb{R}^p : \ell_1 > \ell_2 > \dots > \ell_p > 0\}$. In this paper, these tuples are represented as columns. The sets of all $m \times p$ matrices of real and complex entries are denoted by $\mathbb{R}^{m \times p}$ and $\mathbb{C}^{m \times p}$, respectively. The transpose and the conjugate of \mathbf{C} are denoted by \mathbf{C}'

and $\overline{\mathbf{C}}$, respectively. Furthermore the conjugate transpose of an $m \times p$ matrix $\mathbf{C} \in \mathbb{C}^{m \times p}$ are denoted by $\mathbf{C}^* = \overline{\mathbf{C}}'$. The set of $p \times p$ Hermitian positive definite matrices is denoted by $\mathbb{C}_+^{p \times p}$. For any $\mathbf{c} = \mathbf{a} + \sqrt{-1} \mathbf{b} \in \mathbb{C}^p$ ($\mathbf{a}, \mathbf{b} \in \mathbb{R}^p$), we denote by $[\mathbf{c}]$ a $2p$ -dimensional real vector $(\mathbf{a}', \mathbf{b}')'$. For a positive integer q and real numbers a_1, a_2, \dots, a_q , $\text{Diag}(a_1, a_2, \dots, a_q)$ denotes a $q \times q$ diagonal matrix with the i -th diagonal element a_i ($i = 1, 2, \dots, q$). For an $m \times p$ complex matrix $\mathbf{C} = \mathbf{A} + \sqrt{-1} \mathbf{B}$ ($\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times p}$), we denote by $\{\mathbf{C}\}$ a $2m \times 2p$ real matrix

$$\begin{pmatrix} \mathbf{A} & -\mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}.$$

Let $g(x, y)$ be a real-valued function on an open set $U \in \mathbb{R}^2$. We say that g is differentiable if $\partial g / \partial x$ and $\partial g / \partial y$ exist on U . Let u, v be real-valued functions on an open set $U \in \mathbb{R}^2$. A function $g := u + \sqrt{-1} v$ is called differentiable if u, v are differentiable. For $z = x + \sqrt{-1} y$ ($x, y \in \mathbb{R}$) and differentiable function $g(z) = u(z) + \sqrt{-1} v(z)$, we define

$$\begin{aligned} \frac{\partial}{\partial z} g &= \frac{1}{2} \left(\frac{\partial}{\partial x} - \sqrt{-1} \frac{\partial}{\partial y} \right) g = \frac{1}{2} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + \frac{\sqrt{-1}}{2} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right), \\ \frac{\partial}{\partial \bar{z}} g &= \frac{1}{2} \left(\frac{\partial}{\partial x} + \sqrt{-1} \frac{\partial}{\partial y} \right) g = \frac{1}{2} \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) + \frac{\sqrt{-1}}{2} \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right). \end{aligned}$$

It is checked directly that

$$\frac{\partial}{\partial z} z = 1, \quad \frac{\partial}{\partial z} \bar{z} = 0, \quad \frac{\partial}{\partial \bar{z}} z = 0, \quad \frac{\partial}{\partial \bar{z}} \bar{z} = 1.$$

If g is differentiable, then

$$\overline{\frac{\partial}{\partial z} g} = \frac{\partial}{\partial \bar{z}} \bar{g}. \quad (3)$$

Let $\mathbf{G} = (g_{ij})_{i=1,2,\dots,m,j=1,2,\dots,p}$ be an $m \times p$ matrix, where g_{ij} 's are complex-valued differentiable functions on $\mathbb{C}^{m \times p}$. For $\mathbf{z} = (z_{ij})_{i=1,2,\dots,m,j=1,2,\dots,p} \in \mathbb{C}^{m \times p}$, we set

$$\nabla_{\mathbf{z}} = \left(\frac{\partial}{\partial z_{ij}} \right)_{i=1,2,\dots,m,j=1,2,\dots,p},$$

and we define

$$\text{Re}(\text{Tr}(\nabla'_{\mathbf{z}} \mathbf{G})) = \text{Tr}(\text{Re}(\nabla'_{\mathbf{z}} \mathbf{G})) = \sum_{j=1}^p \sum_{i=1}^m \left\{ \frac{\partial(\text{Re } g_{ij})}{\partial(\text{Re } z_{ij})} + \frac{\partial(\text{Im } g_{ij})}{\partial(\text{Im } z_{ij})} \right\}.$$

2.2 Complex normal distributions and the Stein identity

Recall that a $p \times 1$ complex random vector Z is said to have a p -variate complex normal distribution with a mean vector $\theta \in \mathbb{C}^p$ and a covariance matrix $\Sigma \in \mathbb{C}_+^{p \times p}$ if the density of Z with respect to Lebesgue measure on \mathbb{C}^p is given as

$$f_Z(\mathbf{z}) = \frac{1}{\pi^p} \text{Det}(\Sigma)^{-1} \exp\{-(\mathbf{z} - \theta)^* \Sigma^{-1} (\mathbf{z} - \theta)\}, \quad \mathbf{z} \in \mathbb{C}^p.$$

We use the notation $Z \sim \mathbb{CN}_p(\theta, \Sigma)$ for this.

Lemma 1. *Let Z be a $p \times 1$ complex random vector having $\mathbb{CN}_p(\theta, \Sigma)$ and let $\mathbf{g} = (g_1, g_2, \dots, g_p) : \mathbb{C}^p \rightarrow \mathbb{C}^p$ be differentiable with*

$$\mathbb{E} \left| \frac{\partial (\text{Re } g_i)}{\partial (\text{Re } z_i)} \right|_{\mathbf{z}=Z} < \infty, \quad \mathbb{E} \left| \frac{\partial (\text{Im } g_i)}{\partial (\text{Im } z_i)} \right|_{\mathbf{z}=Z} < \infty, \quad i = 1, 2, \dots, p.$$

Then we have

$$\mathbb{E}[(Z - \theta)^* \Sigma^{-1} \mathbf{g}(Z) + \mathbf{g}^*(Z) \Sigma^{-1} (Z - \theta)] = \mathbb{E} \left[\sum_{i=1}^p \left\{ \frac{\partial (\text{Re } g_i)}{\partial (\text{Re } z_i)} + \frac{\partial (\text{Im } g_i)}{\partial (\text{Im } z_i)} \right\} \right]_{\mathbf{z}=Z}.$$

Proof. Note that

$$(Z - \theta)^* \Sigma^{-1} \mathbf{g}(Z) + \mathbf{g}^*(Z) \Sigma^{-1} (Z - \theta) = 2[Z - \theta]' \{\Sigma^{-1}\} [\mathbf{g}]$$

and that $Z \sim \mathbb{CN}_p(\theta, \Sigma)$ if and only if $[Z] \sim N_{2p}([\theta], (1/2)\{\Sigma\})$, a $2p$ -variate multivariate real normal distribution with a $2p \times 1$ mean vector $[\theta]$ and a $2p \times 2p$ positive definite covariance matrix $(1/2)\{\Sigma\}$. By the Stein identity on a multivariate real normal distribution[see Stein [36]], we have

$$\mathbb{E}\{[Z - \theta]' \{(1/2)\Sigma\}^{-1} [\mathbf{g}]\} = \mathbb{E} \left\{ \sum_{i=1}^p \left\{ \frac{\partial (\text{Re } g_i)}{\partial (\text{Re } z_i)} + \frac{\partial (\text{Im } g_i)}{\partial (\text{Im } z_i)} \right\} \right\}_{\mathbf{z}=Z},$$

which completes the proof. \square

2.3 Complex Wishart distributions and the Stein-Haff identity

Assume that a $p \times p$ Hermitian positive definite matrix \mathbf{S} has a complex Wishart distribution $\text{CW}_p(\Sigma, n)$ with the density function (2). Let $\mathbf{G}(\mathbf{S})$ be a $p \times p$ matrix, the (i, j) element $g_{ij}(\mathbf{S})$

of which is a complex-valued function of $\mathbf{S} = (s_{ij})$. For a $p \times p$ Hermitian matrix $\mathbf{S} = (s_{jk})$, let $\mathbf{D}_S = (\partial/\partial s_{jk})$ be a $p \times p$ operator matrix, the (j, k) element of which is given by

$$\frac{\partial}{\partial s_{jk}} = \frac{1}{2}(1 + \delta_{jk}) \left\{ \frac{\partial}{\partial(\operatorname{Re} s_{jk})} + (1 - \delta_{jk})\sqrt{-1} \frac{\partial}{\partial(\operatorname{Im} s_{jk})} \right\}, \quad j, k = 1, 2, \dots, p. \quad (4)$$

Here δ_{jk} is the Kronecker delta ($= 1$ if $j = k$ and $= 0$ if $j \neq k$). Thus the (j, k) element of $\mathbf{D}_S \mathbf{G}(\mathbf{S})$ is

$$\{\mathbf{D}_S \mathbf{G}(\mathbf{S})\}_{jk} = \sum_{l=1}^p \frac{\partial g_{lk}}{\partial s_{jl}}(\mathbf{S}) = \frac{1}{2}(1 + \delta_{jl}) \sum_{l=1}^p \left\{ \frac{\partial g_{lk}}{\partial(\operatorname{Re} s_{jl})}(\mathbf{S}) + (1 - \delta_{jl})\sqrt{-1} \frac{\partial g_{lk}}{\partial(\operatorname{Im} s_{jl})}(\mathbf{S}) \right\}.$$

It is directly checked that $\partial s_{k\ell}/\partial s_{ij} = \delta_{i\ell}\delta_{jk}$, and that $\partial \bar{s}_{k\ell}/\partial s_{ij} = \delta_{ik}\delta_{j\ell}$.

Lemma 2. *Assume that each entry of $\mathbf{G}(\mathbf{S})$ is a partially differentiable function with respect to $\operatorname{Re} s_{jk}$ and $\operatorname{Im} s_{jk}$, $j, k = 1, 2, \dots, p$. Under conditions on $\mathbf{G}(\mathbf{S})$ specified in Konno [23], the following identity holds:*

$$\mathbb{E}[\operatorname{Tr}(\mathbf{G}(\mathbf{S})\mathbf{\Sigma}^{-1})] = \mathbb{E}[(n - p)\operatorname{Tr}(\mathbf{G}(\mathbf{S})\mathbf{S}^{-1}) + \operatorname{Tr}(\mathbf{D}_S \mathbf{G}(\mathbf{S}))]. \quad (5)$$

Remark 1. The Stein-Haff identity was extended to an elliptically contoured complex distribution by Konno [23]. Hence, if we know the improved estimators for the normal case, we can establish the robustness of improvement for the elliptically contoured complex distribution in a manner similar to that demonstrated in Kubokawa and Srivastava [24, 25].

3 Known covariance case

Estimation of a mean matrix of a real multivariate normal distribution is considered in Stein [34], Efron and Morris [12], Zheng [39, 40], and Ghosh and Sheih [13]. Recently Beran [2, 3] developed adaptive total shrinkage estimators with smaller asymptotic risk than the data matrix.

In this section, we consider a complex analogue of this problem since the known covariance case gives an insight into estimation problem of the mean matrix with unknown covariance matrix. The problem treated in this section is stated as follows: Assume that $m \geq p$ and that we observe an $m \times p$ random matrix \mathbf{Z} with the coordinates z_{ij} ($i = 1, 2, \dots, m, j =$

$1, 2, \dots, p)$ that are independently and identically distributed as $\mathbb{CN}(\xi_{ij}, 1)$ ($\xi_{ij} \in \mathbb{C}$). Set $\Xi = (\xi_{ij})$, i.e., the (i, j) element of an $m \times p$ complex matrix Ξ is given by ξ_{ij} . We use notation $\mathbf{Z} \sim \mathbb{CN}_{m \times p}(\Xi, \mathbf{I}_m \otimes \mathbf{I}_p)$ to indicate that a random matrix \mathbf{Z} has a multivariate complex normal distribution with a mean matrix Ξ and a covariance matrix $\mathbf{I}_m \otimes \mathbf{I}_p$. We consider the problem of estimating the mean matrix Ξ under a loss function

$$\mathcal{L}_0(\hat{\Xi}, \Xi) = \text{Tr}\{(\hat{\Xi} - \Xi)^*(\hat{\Xi} - \Xi)\},$$

where $\hat{\Xi}$ is an estimator of Ξ based on \mathbf{Z} . The risk function is given by

$$\mathcal{R}_0(\hat{\Xi}, \Xi) = \mathbb{E}[\text{Tr}\{(\hat{\Xi} - \Xi)^*(\hat{\Xi} - \Xi)\}],$$

where the expectation is taken with respect to the distribution $\mathbb{CN}_{m \times p}(\Xi, \mathbf{I}_m \otimes \mathbf{I}_p)$.

3.1 Unbiased risk estimate for a class of invariant estimators

The maximum likelihood estimator of Ξ is given by $\hat{\Xi}_{mle} = \mathbf{Z}$ whose risk function is given by $\mathcal{R}_0(\hat{\Xi}_{mle}, \Xi) = mp$. However, it is expected that the estimator $\hat{\Xi}_{mle}$ is improved by so-called shrinkage estimators. In order to search for shrinkage estimators in a systematic way, we introduce the following class of estimators and obtain an unbiased risk estimate for this class. This unbiased risk estimate enables us to find a variety of improved estimators.

Let $\mathbf{W} = \mathbf{Z}^* \mathbf{Z}$ and decompose $\mathbf{W} = \mathbf{U} \mathbf{L} \mathbf{U}^*$, where \mathbf{U} is a $p \times p$ unitary matrix such that $\mathbf{U} \mathbf{U}^* = \mathbf{I}_p$ and $\mathbf{L} = \text{Diag}(\ell_1, \ell_2, \dots, \ell_p)$, a diagonal real matrix whose i -th element ($i = 1, 2, \dots, p$) is given by ℓ_i in the decreasing order. Note that all ℓ_i 's are non-zero with probability one. We consider a class of estimators of the form

$$\hat{\Xi}_H := \hat{\Xi}_H(\mathbf{Z}) = \mathbf{Z}[\mathbf{I}_p + \mathbf{U} \mathbf{H}(\mathbf{L}) \mathbf{U}^*], \quad (6)$$

where $\mathbf{H} := \mathbf{H}(\mathbf{L}) = \text{Diag}(h_1(\mathbf{L}), h_2(\mathbf{L}), \dots, h_p(\mathbf{L}))$ with $h_i(\mathbf{L})$'s, $i = 1, 2, \dots, p$, being real-valued functions on $\mathbb{R}_{>}^p$. This class is a complex version of a class of estimators appeared in Stein [34] and Zheng [39, 40]. The following lemma is the complex counterpart of an unbiased risk estimate for orthogonally invariant class of estimators of a mean matrix of the multivariate real normal distribution, which was proved by Stein [34].

Lemma 3. Assume that $\mathbf{Z} \sim \mathbb{C}N_{m \times p}(\mathbf{\Xi}, \mathbf{I}_m \otimes \mathbf{I}_p)$. For the estimator $\widehat{\mathbf{\Xi}}_H$ given by (6), we have

$$\mathcal{R}_0(\widehat{\mathbf{\Xi}}_H, \mathbf{\Xi}) = mp + \mathbb{E} \left[\sum_{k=1}^p \left\{ 2(m-p+1)h_k(\mathbf{L}) + 2\ell_k h_{kk} + 4 \sum_{b>k} \frac{\ell_k h_k(\mathbf{L}) - \ell_b h_b(\mathbf{L})}{\ell_k - \ell_b} + \ell_k h_k^2(\mathbf{L}) \right\} \right],$$

where $h_{kk} = (\partial h_k / \partial \ell_k)(\mathbf{L})$ ($k = 1, 2, \dots, p$).

Proof. Using Lemma 1 and (6) we have

$$\begin{aligned} \mathcal{R}_0(\widehat{\mathbf{\Xi}}_H, \mathbf{\Xi}) &= \mathbb{E}[\text{Tr} \{ (\mathbf{Z} - \mathbf{\Xi})^* (\mathbf{Z} - \mathbf{\Xi}) + 2 \text{Re} (\nabla'_Z \mathbf{Z} \mathbf{U} \mathbf{H} \mathbf{U}^*) + \mathbf{Z}^* \mathbf{U} \mathbf{H}^2 \mathbf{U}^* \mathbf{Z} \}] \\ &= \mathbb{E} \left[mp + 2 \text{Tr} \{ \text{Re} (\nabla'_Z \mathbf{Z} \mathbf{U} \mathbf{H} \mathbf{U}^*) \} + \sum_{i=1}^p \ell_i h_i^2 \right]. \end{aligned}$$

Use Lemma 6 in the Appendix to evaluate the second term inside expectation of the right hand side of the above equation. \square

Remark 2. We consider the real version of estimating the mean matrix of the multivariate normal distributions. Let $\mathbf{X} \sim N_{m \times p}(\mathbf{\Xi}, \mathbf{I}_m \otimes \mathbf{I}_p)$ and decompose $\mathbf{X}'\mathbf{X} = \mathbf{O}\mathbf{L}\mathbf{O}'$ where \mathbf{O} is a $p \times p$ orthogonal matrix and $\mathbf{L} = \text{Diag}(\ell_1, \ell, \dots, \ell_p)$ are the ordered eigenvalues of $\mathbf{X}'\mathbf{X}$ in decreasing order. Then the unbiased risk estimate for estimators $\widehat{\mathbf{\Xi}}_H = \mathbf{X}[\mathbf{I}_p + \mathbf{O}\mathbf{H}(\mathbf{L})\mathbf{O}']$ is given by

$$\begin{aligned} \mathcal{R}_0(\widehat{\mathbf{\Xi}}_H, \mathbf{\Xi}) &:= \mathbb{E}[\text{Tr} (\mathbf{X} - \mathbf{\Xi})' (\mathbf{X} - \mathbf{\Xi})] \\ &= \mathbb{E} \left[2(m-p+1)h_k(\mathbf{L}) + 4\ell_k h_{kk} + 4 \sum_{b>k} \frac{\ell_k h_k(\mathbf{L}) - \ell_b h_b(\mathbf{L})}{\ell_k - \ell_b} + \ell_k h_k^2(\mathbf{L}) \right], \end{aligned}$$

which can be obtained by replacing the factor 2 of $h_k(\partial h_k / \partial \ell_k)$ in Lemma 3 with 4.

3.2 Alternative estimators

The following proposition is a complex analogue of the results of Zheng [39].

Proposition 1. Assume that $m > p$ and let $\gamma_1(\mathbf{L}), \gamma_2(\mathbf{L}), \dots, \gamma_p(\mathbf{L})$ be functions satisfying

- (i) $0 \leq \gamma_k(\mathbf{L}) \leq 2(m-p)$;
- (ii) $(\partial \gamma_k / \partial \ell_k)(\mathbf{L}) \geq 0$ for $k = 1, 2, \dots, p$;
- (iii) $\gamma_1(\mathbf{L}) \geq \gamma_2(\mathbf{L}) \geq \dots \geq \gamma_p(\mathbf{L})$.

Then the estimator (6) with

$$\mathbf{H} := \mathbf{H}(\mathbf{L}) = -\text{Diag} \left(\frac{\gamma_1(\mathbf{L})}{\ell_1}, \frac{\gamma_2(\mathbf{L})}{\ell_2}, \dots, \frac{\gamma_p(\mathbf{L})}{\ell_p} \right)$$

is minimax.

Proof. From Assumptions (i)–(iii) and Lemma 3, it is easy to show that $\mathcal{R}_0(\mathbf{Z}[\mathbf{I}_p + \mathbf{U}\mathbf{H}\mathbf{U}^*], \mathbf{\Xi}) \leq mp$. \square

Remark 3. Assume that $m > p$. From Proposition 1, it is easily seen that a complex analogue of the crude Efron-Morris estimator $\mathbf{Z}[\mathbf{I}_p - (m - p)(\mathbf{Z}^*\mathbf{Z})^{-1}]$ is minimax.

Proposition 2. Assume that $m > p$ and let $h_k(\mathbf{L}) = -(m + p - 2k)/\ell_k$ ($k = 1, 2, \dots, p$) in (6). Then the estimator of the form (6) is minimax.

Proof. Without loss of generality we can assume that $\ell_1 > \ell_2 > \dots > \ell_p > 0$. Let $h_k(\mathbf{L}) = -c_k/\ell_k$ ($k = 1, 2, \dots, p$) in (6), where c_k 's are positive constants such that $c_1 \geq c_2 \geq \dots \geq c_p$. Then using Lemma 3 and the fact that $\ell_k/(\ell_k - \ell_b) > 1$ for $b > k$, we can see that the risk difference between \mathbf{Z} and $\mathbf{Z}[\mathbf{I}_p + \mathbf{U}\mathbf{H}\mathbf{U}^*]$ is evaluated as

$$\begin{aligned} \Upsilon &= \mathcal{R}_0(\mathbf{Z}, \mathbf{\Xi}) - \mathcal{R}_0(\mathbf{Z}[\mathbf{I}_p + \mathbf{U}\mathbf{H}\mathbf{U}^*], \mathbf{\Xi}) \\ &= \sum_{k=1}^p \mathbb{E} \left[2(m - p) \frac{c_k}{\ell_k} + 4 \sum_{b>k} \frac{c_k - c_b}{\ell_k - \ell_b} - \frac{c_k^2}{\ell_k} \right] \geq \sum_{k=1}^p \mathbb{E} \left[\frac{w_k(c_k)}{\ell_k} \right], \end{aligned}$$

where

$$w_k(t) = 2(m + p - 2k)t - t^2 - 4 \sum_{b=k+1}^p c_b.$$

If $c_b = m + p - 2b$ ($b = k + 1, \dots, p$), then each $w_k(t)$ is maximized at $t = (m + p - 2k)$. Hence, for $c_k = m + p - 2k$ ($k = 1, 2, \dots, p$), we can see that $w_k(m + p - 2k) = w_{k-1}(m + p - 2k) < w_{k-1}(m + p - 2(k - 1))$ for $k = 2, \dots, p$. Therefore we have

$$0 < w_p(m - p) < w_{p-1}(m - p + 2) < \dots < w_2(m + p - 4) < w_1(m + p - 2),$$

from which it follows that $\Upsilon > 0$. \square

Remark 4. It is easy to extend the result to a known correlated covariance case. Assume that we observe an $m \times p$ random matrix $\tilde{\mathbf{Z}}$ that is distributed as $\mathbb{C}N_{m \times p}(\tilde{\mathbf{\Xi}}, \mathbf{I}_m \otimes \mathbf{\Sigma})$ with an $m \times p$ unknown complex matrix $\tilde{\mathbf{\Xi}}$ and a known $p \times p$ positive definite Hermitian matrix $\mathbf{\Sigma}$. Consider the problem of estimating $\tilde{\mathbf{\Xi}}$ under the loss function $\text{Tr}\{(\hat{\tilde{\mathbf{\Xi}}} - \tilde{\mathbf{\Xi}})^*(\hat{\tilde{\mathbf{\Xi}}} - \tilde{\mathbf{\Xi}})\mathbf{\Sigma}^{-1}\}$, where $\hat{\tilde{\mathbf{\Xi}}}$ is an estimator of $\tilde{\mathbf{\Xi}}$. Transforming $\tilde{\mathbf{Z}} \rightarrow \tilde{\mathbf{Z}}\mathbf{\Sigma}^{-1/2} =: \mathbf{Z}$, $\tilde{\mathbf{\Xi}} \rightarrow \tilde{\mathbf{\Xi}}\mathbf{\Sigma}^{-1/2} =: \mathbf{\Xi}$, and $\hat{\tilde{\mathbf{\Xi}}} \rightarrow \hat{\tilde{\mathbf{\Xi}}}\mathbf{\Sigma}^{-1/2} =: \hat{\mathbf{\Xi}}$, the problem reduces to the case when $\mathbf{\Sigma} = \mathbf{I}_p$. Therefore, the Efron-Morris estimator of $\tilde{\mathbf{\Xi}}$ is given by $\tilde{\mathbf{Z}}[\mathbf{I}_p - (m - p)(\tilde{\mathbf{Z}}^*\tilde{\mathbf{Z}})^{-1}\mathbf{\Sigma}]$. If $\mathbf{\Sigma}$ is unknown and if we observe $\mathbf{S} \sim \mathbb{C}W_p(\mathbf{\Sigma}, n)$, we replace $\mathbf{\Sigma}$ with \mathbf{S}/n to obtain an estimator $\tilde{\mathbf{Z}}[\mathbf{I}_p - ((m - p)/n)(\tilde{\mathbf{Z}}^*\tilde{\mathbf{Z}})^{-1}\mathbf{S}]$. This form of estimators is developed in Section 4. Similarly it is easily seen that the Efron-Morris estimator of $\tilde{\mathbf{\Xi}}$ is given by $[\mathbf{I}_m - (p - m)(\tilde{\mathbf{Z}}\mathbf{\Sigma}^{-1}\tilde{\mathbf{Z}}^*)^{-1}]\tilde{\mathbf{Z}}$ if $p > m$. If $\mathbf{\Sigma}$ is unknown and if we observe $\mathbf{S} \sim \mathbb{C}W_p(\mathbf{\Sigma}, n)$, we replace $\mathbf{\Sigma}$ with \mathbf{S}/n to obtain an estimator $[\mathbf{I}_m - ((p - m)/n)(\tilde{\mathbf{Z}}\mathbf{S}^{-1}\tilde{\mathbf{Z}}^*)^{-1}]\tilde{\mathbf{Z}}$. This form of estimators is also developed in Section 4.

4 Unknown case and invariant loss

In the sequel of this section and the Appendix, we assume that $\mathbf{K} = \mathbf{I}_m$ since the results for the known matrix \mathbf{K} can be obtained from those for $\mathbf{K} = \mathbf{I}_m$ by using a transformation similar to that given in Remark 4.

4.1 Unbiased risk estimate for a class of invariant estimators

Consider a class of estimators of the form $\mathbf{Z} + \mathbf{G}(\mathbf{Z}, \mathbf{S})$, where $\mathbf{G} := \mathbf{G}(\mathbf{Z}, \mathbf{S})$ is an $m \times p$ matrix whose (i, j) element g_{ij} ($i = 1, 2, \dots, m, j = 1, 2, \dots, p$) is a complex-valued function based on (\mathbf{Z}, \mathbf{S}) .

Lemma 4. *Assume that all elements of $\mathbf{G}(\mathbf{Z}, \mathbf{S})$ are absolutely continuous functions of \mathbf{Z} and \mathbf{S} . Then we have*

$$\begin{aligned} \mathcal{R}(\mathbf{Z} + \mathbf{G}(\mathbf{Z}, \mathbf{S}), (\mathbf{\Xi}, \mathbf{\Sigma})) &= mp + \mathbb{E}[2 \text{Tr}\{\text{Re}(\nabla'_Z \mathbf{G}(\mathbf{Z}, \mathbf{S}))\} + \text{Tr}\{\mathbf{D}_S \mathbf{G}^*(\mathbf{Z}, \mathbf{S})\mathbf{G}(\mathbf{Z}, \mathbf{S})\} \\ &\quad + (n - p)\text{Tr}\{\mathbf{G}^*(\mathbf{Z}, \mathbf{S})\mathbf{G}(\mathbf{Z}, \mathbf{S})\mathbf{S}^{-1}\}]. \end{aligned} \quad (7)$$

Proof. Use Lemmas 1 and 2. \square

To describe our class of estimators, let $\mathbf{F} = \text{Diag}(f_1, f_2, \dots, f_{\min(m,p)})$ be the eigenvalues of $\mathbf{Z}^*\mathbf{Z}\mathbf{S}^{-1}$. For $p > m$ decompose $\mathbf{Z}\mathbf{S}^{-1}\mathbf{Z}^* = \mathbf{U}\mathbf{F}\mathbf{U}^*$, where \mathbf{U} is an $m \times m$ unitary matrix. For $m > p$ we decompose $\mathbf{S} = (\mathbf{A}^*)^{-1}\mathbf{A}^{-1}$ and $\mathbf{Z}^*\mathbf{Z} = (\mathbf{A}^*)^{-1}\mathbf{F}\mathbf{A}^{-1}$, where \mathbf{A} is a $p \times p$ non-singular matrix. We consider a class of estimators of the form

$$\widehat{\Xi}_H := \widehat{\Xi}_H(\mathbf{Z}, \mathbf{S}) = \begin{cases} \mathbf{Z}\{\mathbf{I}_p + \mathbf{A}\mathbf{H}(\mathbf{F})\mathbf{A}^{-1}\} & \text{if } m > p \\ \{\mathbf{I}_m + \mathbf{U}\mathbf{H}(\mathbf{F})\mathbf{U}^*\}\mathbf{Z} & \text{if } p > m \end{cases}, \quad (8)$$

where $\mathbf{H} := \mathbf{H}(\mathbf{F}) = \text{Diag}(h_1(\mathbf{F}), h_2(\mathbf{F}), \dots, h_{\min(m,p)}(\mathbf{F}))$ whose i -th element $h_i := h_i(\mathbf{F})$, $i = 1, 2, \dots, \min(m, p)$, is a real-valued function on $\mathbb{R}_{>}^{\min(m,p)}$.

Let

$$\begin{aligned} \widehat{\Delta}(n, m, p; \mathbf{H}) &= \sum_{k=1}^p \left\{ 2(m-p+1)h_k(\mathbf{F}) + 2f_k h_{kk}(\mathbf{F}) + 4 \sum_{b>k} \frac{f_k h_k(\mathbf{F}) - f_b h_b(\mathbf{F})}{f_k - f_b} \right. \\ &\quad \left. + (n+p-2)f_k h_k^2(\mathbf{F}) - 2f_k^2 h_{kk}(\mathbf{F}) h_k(\mathbf{F}) - 2 \sum_{b>k} \frac{f_k^2 h_k^2(\mathbf{F}) - f_b^2 h_b^2(\mathbf{F})}{f_k - f_b} \right\}, \quad (9) \end{aligned}$$

where $h_{kk}(\mathbf{F}) = (\partial h_k / \partial f_k)(\mathbf{F})$, $k = 1, 2, \dots, p$.

Proposition 3. *Under the suitable conditions, we have*

$$\mathcal{R}(\widehat{\Xi}_H, (\Xi, \Sigma)) = \begin{cases} mp + \mathbb{E}[\widehat{\Delta}(n, m, p; \mathbf{H})] & \text{if } m > p \\ mp + \mathbb{E}[\widehat{\Delta}(n+m-p, p, m; \mathbf{H})] & \text{if } p > m \end{cases}.$$

Proof. We apply Lemmas 8 and 10 in the Appendix to (7). For $m > p$, set $\Phi = \mathbf{H}$ in the first equation of Lemma 8 and $\Phi = \mathbf{F}\mathbf{H}^2$ in the second equation of Lemma 8 to get the risk representation for the case when $m > p$. For $p > m$, set $\Phi = \mathbf{H}$ in the first equation of Lemma 10 and $\Phi = \mathbf{H}^2$ in the second equation of Lemma 10 to conclude the proof of the risk representation for the case when $p > m$. \square

Remark 5. Assume that $\Xi = \mathbf{0}$ in (1). From [20], the joint distribution of the eigenvalues of $\mathbf{Z}^*\mathbf{Z}\mathbf{S}^{-1}$ is, aparting from normalizing constants,

$$\prod_{k=1}^p \frac{f_k^{m-p}}{(1+f_k)^{n+m}} \prod_{k=1}^{p-1} \prod_{j=k+1}^p (f_k - f_j)^2 \prod_{k=1}^p df_k$$

if $m > p$ while it is

$$\prod_{k=1}^m \frac{f_k^{p-m}}{(1+f_k)^{n+m}} \prod_{k=1}^{m-1} \prod_{j=k+1}^m (f_k - f_j)^2 \prod_{k=1}^m df_k$$

if $p > m$. Note that the substitution rule to get the second distribution from the first distribution, i.e.,

$$(p, m, n) \rightarrow (m, p, n + m - p)$$

is valid to obtain the second assertion of Proposition 3 from the first assertion of Proposition 3. Hence, if we know the estimator of the form $\mathbf{Z}\{\mathbf{I}_p + \mathbf{A}\mathbf{H}\mathbf{A}^{-1}\}$ when $m > p$, we can easily write down estimators of the form $\{\mathbf{I}_m + \mathbf{U}\mathbf{H}\mathbf{U}^*\}\mathbf{Z}$ when $p > m$ by using the above substitution rule.

Remark 6. We consider the real version of estimating the mean matrix of the multivariate normal distributions. Let $\mathbf{X} \sim N_{m \times p}(\boldsymbol{\Xi}, \mathbf{I}_m \otimes \boldsymbol{\Sigma}_p)$ and $\mathbf{S} \sim W_p(n, \boldsymbol{\Sigma})$, where \mathbf{X} and \mathbf{S} are independent. Let $\mathbf{F} = \text{Diag}(f_1, f_2, \dots, f_{\min(m,p)})$ be the eigenvalues of $\mathbf{X}'\mathbf{X}\mathbf{S}^{-1}$. For $p > m$ decompose $\mathbf{X}\mathbf{S}^{-1}\mathbf{X}' = \mathbf{O}\mathbf{F}\mathbf{O}'$, where \mathbf{O} is an $m \times m$ orthogonal matrix. For $m > p$ we decompose $\mathbf{S} = (\mathbf{A}')^{-1}\mathbf{A}^{-1}$ and $\mathbf{X}'\mathbf{X} = (\mathbf{A}')^{-1}\mathbf{F}\mathbf{A}^{-1}$, where \mathbf{A} is a $p \times p$ non-singular matrix. We consider a class of estimators of the form $\hat{\boldsymbol{\Xi}}_H := \mathbf{X}\{\mathbf{I}_p + \mathbf{A}\mathbf{H}(\mathbf{F})\mathbf{A}^{-1}\}$ where $m > p$, and $\mathbf{H} := \mathbf{H}(\mathbf{F}) = \text{Diag}(h_1(\mathbf{F}), h_2(\mathbf{F}), \dots, h_{\min(m,p)}(\mathbf{F}))$ whose i -th element $h_i := h_i(\mathbf{F})$, $i = 1, 2, \dots, \min(m, p)$, is a real-valued function on $\mathbb{R}_{>}^{\min(m,p)}$. Then the real version of unbiased estimate for the class of estimators $\hat{\boldsymbol{\Xi}}_H$ is obtained by changing the coefficients of terms $f_k h_{kk}$, $f_k h_k^2$, and $f_k h_{kk} h_k$. The real version of $\hat{\Delta}(n, m, p; \mathbf{H})$ in (9) is given as

$$\sum_{k=1}^p \left\{ 2(m-p+1)h_k(\mathbf{F}) + 4f_k h_{kk}(\mathbf{F}) + 4 \sum_{b>k} \frac{f_k h_k(\mathbf{F}) - f_b h_b(\mathbf{F})}{f_k - f_b} + (n+p-3)f_k h_k^2(\mathbf{F}) - 4f_k^2 h_{kk}(\mathbf{F}) h_k(\mathbf{F}) - 2 \sum_{b>k} \frac{f_k^2 h_k^2(\mathbf{F}) - f_b^2 h_b^2(\mathbf{F})}{f_k - f_b} \right\}.$$

4.2 Alternative estimators

Proposition 4. Let $\gamma_1(\mathbf{F}), \gamma_2(\mathbf{F}), \dots, \gamma_{\min(m,p)}(\mathbf{F})$ be functions satisfying

- (i) $0 \leq \gamma_k(\mathbf{F}) \leq \max\{2(m-p)/(n+p), 2(p-m)/(n+2m-p)\}$;
- (ii) $(\partial\gamma_k/\partial f_k)(\mathbf{F}) \geq 0$ for $k = 1, 2, \dots, \min(m, p)$;
- (iii) $\gamma_1(\mathbf{F}) \geq \gamma_2(\mathbf{F}) \geq \dots \geq \gamma_{\min(m,p)}(\mathbf{F})$.

Then the estimator (8) with

$$\mathbf{H}(\mathbf{F}) = -\text{Diag} \left(\frac{\gamma_1(\mathbf{F})}{f_1}, \frac{\gamma_2(\mathbf{F})}{f_2}, \dots, \frac{\gamma_{\min(m,p)}(\mathbf{F})}{f_{\min(m,p)}} \right)$$

is minimax.

Proof. From Assumptions (i)–(iii) and Proposition 3, it is easy to check that $\mathcal{R}(\widehat{\boldsymbol{\Xi}}_H, (\boldsymbol{\Xi}, \boldsymbol{\Sigma})) \leq mp$.

□

Corollary 1. *The Efron-Morris estimator*

$$\widehat{\boldsymbol{\Xi}}^{(EM)} = \begin{cases} \mathbf{Z} \left\{ \mathbf{I}_p - \frac{m-p}{n+p} (\mathbf{Z}^* \mathbf{Z})^{-1} \mathbf{S} \right\} & \text{if } m > p \\ \left\{ \mathbf{I}_m - \frac{p-m}{n+2m-p} (\mathbf{Z} \mathbf{S}^{-1} \mathbf{Z}^*)^{-1} \right\} \mathbf{Z} & \text{if } p > m \end{cases}$$

is minimax.

Proof. It is immediately seen from Proposition 4.

□

Proposition 5. *For $k = 1, 2, \dots, \min(m, p)$, let*

$$c_k^{(AS)} = \frac{m+p-2k}{n-p+2k}, \quad \mathbf{H}^{(AS)}(\mathbf{F}) = -\text{Diag} \left(\frac{c_1^{(AS)}}{f_1}, \frac{c_2^{(AS)}}{f_2}, \dots, \frac{c_{\min(m,p)}^{(AS)}}{f_{\min(m,p)}} \right).$$

Then the estimator

$$\widehat{\boldsymbol{\Xi}}^{(AS)} = \begin{cases} \mathbf{Z} \left\{ \mathbf{I}_p + \mathbf{A} \mathbf{H}^{(AS)}(\mathbf{F}) \mathbf{A}^{-1} \right\} & \text{if } m > p \\ \left\{ \mathbf{I}_m + \mathbf{U} \mathbf{H}^{(AS)}(\mathbf{F}) \mathbf{U}^* \right\} \mathbf{Z} & \text{if } p > m \end{cases}$$

is minimax.

Proof. It suffices to prove that $\widehat{\Delta} := \widehat{\Delta}(n, m, p; \mathbf{H}) \leq 0$ for the case when $f_1 > f_2 > \dots > f_p > 0$ and $m > p$. Put $h_k = -c_k/f_k$ for $k = 1, 2, \dots, p$ in (9), where c_k 's are positive constants such that $c_1 \geq c_2 \geq \dots \geq c_p$. Note that

$$\begin{aligned} \sum_{k=1} \sum_{b>k} \left\{ \frac{4(c_k - c_b)}{f_k - f_b} + \frac{2(c_k^2 - c_b^2)}{f_k - f_b} \right\} &= \sum_{k=1} \sum_{b>k} \left\{ \frac{2(c_k - c_b)(2 + c_k + c_b)}{f_k - f_b} \right\} \\ &\geq \sum_{k=1} \sum_{b>k} \left\{ \frac{2(c_k - c_b)(2 + c_k + c_b)}{f_k} \right\} \\ &= \sum_{k=1} \frac{1}{f_k} \left\{ 4(p-k)c_k + 2(p-k)c_k^2 - \sum_{b>k} \{4c_b + 2c_b^2\} \right\}. \end{aligned}$$

Hence, for $\mathbf{H} = -\text{Diag}(c_1/f_1, c_2/f_2, \dots, c_p/f_p)$, we have

$$\hat{\Delta} \leq -\sum_{k=1}^p \frac{1}{f_k} \left\{ 2(m+p-2k)c_k - (n-p+2k)c_k^2 - 2 \sum_{b>k} \{2c_b + c_b^2\} \right\}.$$

Proceed in a way similar to the proof of Proposition 2 to see that the right hand side of the above inequality is negative. \square

Acknowledgements

This work was in part supported by the Japan Society for the Promotion of Science through Grants-in-Aid for Scientific Research (C) (No.21500283).

References

- [1] H.H. Andersen, M. Højbjerg, D. Sørensen, P.S. Eriksen, **LINEAR AND GRAPHICAL MODELS**, Springer-Verlag, New York (1995).
- [2] R. Beran, Adaptive estimators of a mean matrix: Total least squares versus total shrinkage, **ECONOMETRIC THEORY** **24** (2008) 448–471.
- [3] R. Beran, Estimating a mean matrix: Boosting efficiency by multiple affine shrinkage, **ANN. INST. STATIST. MATH.** **60** (2008) 843–864.
- [4] M. Bilodeau, Multivariate flattening for better predictions, **CAN. J. STATIST** **28** (2000) 159–170.
- [5] M. Bilodeau, T. Kariya, Minimax estimators in the normal MANOVA model, **J. MULTIVARIATE ANAL.** **28** (1989) 260–270.
- [6] L. Breiman, J.H. Friedman, Predicting multivariate responses in multiple regression, **J. ROY. STATIST. SOC. SER. B** **59** (1998) 3–54.
- [7] P.J. Brown, J.V. Zidek, Adaptive multivariate ridge regression, **ANN. STATIST.** **8** (1980) 64–74.

- [8] P.J. Brown, J.V. Zidek, Multivariate regression shrinkage estimators with unknown covariance matrix, *SCAND. J. STATIST.* **9** (1982) 209–215.
- [9] A.P. Dempster, M. Schatoff, N. Wermuth, A simulation study of alternative to ordinary least squares, *J. AMER. STAT. ASSOC.* **72** (1977) 77–91.
- [10] J.A. Díaz-García, R. Gutierréz-Jáimez. Matricvariate and matrix multivariate T distributions and associated distributions, *METRIKA* **75** (2012) 963-976.
- [11] A. DoGondzić, A. Neborai, Generalized multivariate analysis of variance, *IEEE SIGNAL PROCESSING MAGAZIN*, September (2003) 39-54.
- [12] B. Efron, C. Morris, Families of minimax estimators of the mean of a multivariate normal distribution, *ANN. STATIST.* **4** (1976) 11–21.
- [13] M. Ghosh, G. Shieh, Empirical Bayes minimax estimators of matrix normal means, *J. MULTIVARIATE ANAL.* **38** (1991) 306–318.
- [14] N.R. Goodman, Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction), *ANN. MATH. STATIST.* **34** (1963) 152-176.
- [15] L.R. Haff, An identity for the Wishart distribution with application, *J. MULTIVARIATE ANAL.* **9** (1979) 531–542.
- [16] L.R. Haff, Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity, *ANN. STATIST.* **7** (1979) 1264-1276.
- [17] T. Kariya, Y. Konno, W.E. Strawderman, Double shrinkage estimators in the GMANOVA model, *J. MULTIVARIATE ANAL.* **56** (1996) 245–258.
- [18] T. Kariya, Y. Konno, W.E. Strawderman, Construction of shrinkage estimators for the regression coefficient matrix in GMANOVA model, *COMMUN. STATIST. THEORY METHODS* **28** (1999) 597–611.

- [19] S.M. Kay, FUNDAMENTALS OF STATISTICAL SIGNAL PROCESSING: ESTIMATION THEORY, Prentice Hall PTR (1993).
- [20] C.G. Khatri, Classical statistical analysis based on a certain multivariate complex Gaussian distribution, ANN. MATH. STATIST. **36** (1965) 98-114.
- [21] Y. Konno, On estimation of a matrix of normal means with unknown covariance matrix. J. MULTIVARIATE ANAL. **36** (1991) 44–55.
- [22] Y. Konno, Improved estimation of matrix of normal mean and eigenvalues in the multivariate F-distributions, Doctoral dissertation, University of Tsukuba (1992).
- [23] Y. Konno, Improving on the sample covariance matrix for a complex elliptically contoured distribution. J. STATIST. PLAN. INFER. **137** (2007) 2237-2253.
- [24] T. Kubokawa, M.S. Srivastava, Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution, ANN. STATIST. **27** (1999) 600–609.
- [25] T. Kubokawa, M.S. Srivastava, Robust improvement in estimation of a mean matrix in an elliptically contoured distribution, J. MULTIVARIATE ANAL. **76** (2001) 138–152.
- [26] J. Lillstøl, Improved estimators for multivariate complex-normal regression with application to analysis of linear time-invariant relationships, J. MULTIVARIATE ANAL. **7** (1977) 512–524.
- [27] W.L. Loh, Estimating covariance matrices, Ph.D. thesis, Stanford University (1988).
- [28] W.L. Loh, Estimating covariance matrices, ANN. STATIST. **19** (1991) 283–296.
- [29] W.L. Loh, Estimating covariance matrices II, J. MULTIVARIATE ANAL. **36** (1991) 263–174.
- [30] A.C. Micheas, D.K. Dey, K.V. Mardia, Complex elliptical distributions with application to shape theory, J. STAT. PLAN. INFERENCE **136** (2006) 2961-2982.

- [31] S.D. Oman, Minimax hierarchical empirical Bayes estimation in multivariate regression, *J. MULTIVARIATE ANAL.* **80** (2002) 285–301.
- [32] T. Ratnarajah, R. Vaillancourt, A. Alvo, Complex random matrices and Rician channel capacity. *PROBL. INF. TRANSM.* **41** (2005) 1-22.
- [33] M.S. Srivastava, and T.K.S. Solanky, Predicting multivariate response in linear regression model, *COMM. STATIST. SIMULATION COMPUT.* **32** (2003) 389–409.
- [34] C. Stein, Estimation of the mean of a multivariate normal distribution, IN *PROC. PRAGUE SYMP. ASYMPTOTIC STATIST.* (1974) 345–381.
- [35] C. Stein, Lectures on the theory of estimation of many parameters, in *STUDIES IN THE STATISTICAL THEORY OF ESTIMATION I*(I. A. Ibragimov and M. S. Nikulin, eds.) (1977).
- [36] C. Stein, Estimation of the mean of a multivariate normal distribution, *ANN. STATIST.* **9** (1981) 1135–1151.
- [37] L. Svensson, M. Lundberg, Estimating complex covariance matrix, *SIGNALS, SYSTEMS AND COMPUTERS, CONFERENCE RECORD OF THE THIRTY-EIGHTH ASILOMAR CONFERENCE ON* (2004) 7-10, 2151 - 2154.
- [38] A. van der Merwe, J.V. Zidek, Multivariate regression analysis and canonical variates, *CAN. J. STATIST.* **8** (1980), 27–39.
- [39] Z. Zheng, On estimation of matrix of normal mean, *J. MULTIVARIATE ANAL.* **18** (1986) 70–82.
- [40] Z. Zheng, Selecting a minimax estimator doing well at a point, *J. MULTIVARIATE ANAL.* **19** (1986) 14–23.
- [41] J.V. Zidek, Deriving unbiased risk estimators of multinormal mean and regression coefficient estimators using zonal polynomials, *ANN. STATIST.* **6** (1978) 679–782.

A Appendix

This section develops somewhat tedious computations on eigenvalues, which is a complex analogue of the results given by Loh [27, 28, 29] and Konno [22]. In Section A.1, we give some results which is needed to prove Lemma 3. In Sections A.2 and A.3, we provide some results which is needed to prove Proposition 3.

In the sequel of this section, we use the following notation: For a $p \times q$ real matrix $\mathbf{X} = (x_{ij})_{i=1,2,\dots,m,j=1,2,\dots,p}$, define the matrix of differential as $(d\mathbf{X}) = (dx_{ij})$. We also use the notation $(d\mathbf{X})_{ij} := dx_{ij}$. For a $m \times p$ complex matrix $\mathbf{X} = \mathbf{X}_1 + \sqrt{-1}\mathbf{X}_2$, where $\mathbf{X}_1, \mathbf{X}_2$ are real matrices, we write $(d\mathbf{X}) = (d\mathbf{X}_1) + \sqrt{-1}(d\mathbf{X}_2)$.

A.1 Eigencalculus for known covariance case

Let $\mathbf{W} = \mathbf{Z}^*\mathbf{Z} = \mathbf{U}\mathbf{L}\mathbf{U}^*$, where $\mathbf{U} = (u_{ij})_{i,j=1,\dots,p}$ is a $p \times p$ unitary matrix and $\mathbf{L} = \text{Diag}(\ell_1, \ell_2, \dots, \ell_p)$ with diagonal elements $\ell_1, \ell_2, \dots, \ell_p$ ($\ell_1 > \ell_2 > \dots > \ell_p > 0$). Recall that

$$\frac{\partial}{\partial z_{ij}} = \frac{1}{2} \left(\frac{\partial}{\partial(\text{Re } z_{ij})} - \sqrt{-1} \frac{\partial}{\partial(\text{Im } z_{ij})} \right).$$

Lemma 5. *We have*

$$\begin{aligned} \frac{\partial u_{il}}{\partial z_{jk}} &= \sum_b \sum_{c \neq l} \frac{u_{ic} \bar{u}_{bc} u_{kl} \bar{z}_{jb}}{\ell_l - \ell_c}, \\ \frac{\partial \bar{u}_{il}}{\partial z_{jk}} &= \sum_b \sum_{c \neq l} \frac{\bar{u}_{ic} \bar{u}_{bl} u_{kc} \bar{z}_{jb}}{\ell_l - \ell_c}, \\ \frac{\partial \ell_i}{\partial z_{jk}} &= \sum_b u_{ki} \bar{u}_{bi} \bar{z}_{jb}. \end{aligned}$$

Proof. Taking the differential of $\mathbf{W} = \mathbf{U}\mathbf{L}\mathbf{U}^*$ we obtain that

$$d\mathbf{W} = (d\mathbf{U})\mathbf{L}\mathbf{U}^* + \mathbf{U}\mathbf{L}(d\mathbf{U}^*) + \mathbf{U}(d\mathbf{L})\mathbf{U}^*.$$

Multiplying on the left by \mathbf{U}^* and on the right by \mathbf{U} we have

$$\mathbf{U}^*(d\mathbf{W})\mathbf{U} = (\mathbf{U}^*(d\mathbf{U}))\mathbf{L} + \mathbf{L}(\mathbf{U}^*(d\mathbf{U}))^* + d\mathbf{L}. \quad (10)$$

But, taking the differential of $\mathbf{U}^*\mathbf{U} = \mathbf{I}_p$, we get

$$(\mathrm{d}\mathbf{U}^*)\mathbf{U} + \mathbf{U}^*(\mathrm{d}\mathbf{U}) = \mathbf{0}. \quad (11)$$

Reverting to the coordinates, we obtain from (10) and (11)

$$(\mathbf{U}^*(\mathrm{d}\mathbf{U}))_{ij} = \begin{cases} \frac{1}{\ell_j - \ell_i} (\mathbf{U}^*(\mathrm{d}\mathbf{W})\mathbf{U})_{ij} & \text{for } i \neq j, \\ 0 & \text{for } i = j, \end{cases} \quad (12)$$

and

$$(\mathrm{d}\mathbf{L})_{ii} = (\mathbf{U}^*(\mathrm{d}\mathbf{W})\mathbf{U})_{ii}. \quad (13)$$

We note that

$$(\mathrm{d}\mathbf{U})_{il} = \sum_c u_{ic} (\mathbf{U}^*(\mathrm{d}\mathbf{U}))_{cl} = \sum_{c \neq l} \frac{u_{ic}}{\ell_l - \ell_c} (\mathbf{U}^*(\mathrm{d}\mathbf{W})\mathbf{U})_{cl} = \sum_{b_1, b_2} \sum_{c \neq l} \frac{u_{ic} \bar{u}_{b_1 c} u_{b_2 l}}{\ell_l - \ell_c} (\mathrm{d}\mathbf{W})_{b_1 b_2}.$$

But, from $\mathbf{W} = \mathbf{Z}^*\mathbf{Z}$, we observe that

$$\begin{aligned} (\mathrm{d}\mathbf{W})_{b_1 b_2} \left(\frac{\partial}{\partial z_{jk}} \right) &= \sum_c \left\{ \bar{z}_{cb_1} (\mathrm{d}\mathbf{Z})_{cb_2} \left(\frac{\partial}{\partial z_{jk}} \right) + z_{cb_2} (\mathrm{d}\bar{\mathbf{Z}})_{cb_1} \left(\frac{\partial}{\partial z_{jk}} \right) \right\} = \sum_c \bar{z}_{cb_1} \delta_{jc} \delta_{kb_2} \\ &= \bar{z}_{jb_1} \delta_{kb_2}, \end{aligned} \quad (14)$$

from which it follows that

$$\begin{aligned} \frac{\partial u_{il}}{\partial z_{jk}} &= (\mathrm{d}\mathbf{U})_{il} \left(\frac{\partial}{\partial z_{jk}} \right) \\ &= \sum_{b_1, b_2} \sum_{c \neq l} \frac{u_{ic} \bar{u}_{b_1 c} u_{b_2 l}}{\ell_l - \ell_c} (\mathrm{d}\mathbf{W})_{b_1 b_2} \left(\frac{\partial}{\partial z_{jk}} \right) \\ &= \sum_{b_1, b_2} \sum_{c \neq l} \frac{u_{ic} \bar{u}_{b_1 c} u_{b_2 l} \bar{z}_{jb_1} \delta_{kb_2}}{\ell_l - \ell_c} \\ &= \sum_{b_1} \sum_{c \neq l} \frac{u_{ic} \bar{u}_{b_1 c} u_{kl} \bar{z}_{jb_1}}{\ell_l - \ell_c}. \end{aligned}$$

This completes the first equation of this lemma.

To prove the second equation we take the complex conjugate of (12) and get

$$(\mathbf{U}'(\mathrm{d}\bar{\mathbf{U}}))_{ij} = \begin{cases} \frac{1}{\ell_j - \ell_i} (\mathbf{U}'(\mathrm{d}\bar{\mathbf{W}})\bar{\mathbf{U}})_{ij} & \text{for } i \neq j, \\ 0 & \text{for } i = j. \end{cases}$$

Using the above equation and noting that $\overline{\mathbf{W}} = \mathbf{W}'$ since \mathbf{W} is Hermitian, we have

$$\begin{aligned}
\frac{\partial \bar{u}_{il}}{\partial z_{jk}} &= (\mathrm{d}\bar{\mathbf{U}})_{il} \left(\frac{\partial}{\partial z_{jk}} \right) = \sum_c \bar{u}_{ic} (\mathbf{U}'(\mathrm{d}\bar{\mathbf{U}}))_{cl} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{c \neq l} \frac{\bar{u}_{ic}}{\ell_l - \ell_c} (\mathbf{U}'(\mathrm{d}\bar{\mathbf{W}})\bar{\mathbf{U}})_{cl} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{b_1, b_2} \sum_{c \neq l} \frac{\bar{u}_{ic} u_{b_1 c} \bar{u}_{b_2 l}}{\ell_l - \ell_c} (\mathrm{d}\mathbf{W})_{b_2 b_1} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{b_1, b_2} \sum_{c \neq l} \frac{\bar{u}_{ic} u_{b_1 c} \bar{u}_{b_2 l} \bar{z}_{j b_2} \delta_{k b_1}}{\ell_l - \ell_c} \\
&= \sum_{b_2} \sum_{c \neq l} \frac{\bar{u}_{ic} \bar{u}_{b_2 l} u_{k c} \bar{z}_{j b_2}}{\ell_l - \ell_c},
\end{aligned}$$

which completes the proof of the second equation of this lemma. The third equality follows from the fact that \mathbf{W} is Hermitian while the forth equality follows from (14).

Finally, by (13) and (14), we have

$$\begin{aligned}
\frac{\partial \ell_i}{\partial z_{jk}} &= (\mathrm{d}\mathbf{L})_{ii} \left(\frac{\partial}{\partial z_{jk}} \right) = (\mathbf{U}^*(\mathrm{d}\mathbf{W})\mathbf{U})_{ii} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{b_1, b_2} \bar{u}_{b_1 i} u_{b_2 i} (\mathrm{d}\mathbf{W})_{b_1 b_2} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{b_1, b_2} \bar{u}_{b_1 i} u_{b_2 i} \bar{z}_{j b_1} \delta_{k b_2} = \sum_{b_1} \bar{u}_{b_1 i} u_{k i} \bar{z}_{j b_1},
\end{aligned}$$

which completes the proof of the third equation of this lemma. \square

Lemma 6. Let $\Phi(\mathbf{L}) = \text{Diag}(\varphi_1(\mathbf{L}), \varphi_2(\mathbf{L}), \dots, \varphi_p(\mathbf{L}))$, where $\varphi_i(\mathbf{L})$'s ($i = 1, 2, \dots, p$) are differentiable functions from $\mathbb{R}_{>}^p \rightarrow \mathbb{R}_+$. Then we have

$$\text{Tr} \{ \text{Re}(\nabla'_Z \mathbf{Z} \mathbf{U} \Phi(\mathbf{L}) \mathbf{U}^*) \} = \sum_k \left\{ (m - p + 1) \varphi_k(\mathbf{L}) + 2 \sum_{c > k} \frac{\ell_k \varphi_k(\mathbf{L}) - \ell_c \varphi_c(\mathbf{L})}{\ell_k - \ell_c} + \ell_k \frac{\partial \varphi_k}{\partial \ell_k}(\mathbf{L}) \right\}.$$

Proof. Write Φ and φ_i for $\Phi(\mathbf{L})$ and $\varphi_i(\mathbf{L})$ ($i = 1, 2, \dots, p$), respectively. Note that

$$\text{Tr} \{ \text{Re}(\nabla'_Z \mathbf{Z} \mathbf{U} \Phi \mathbf{U}^*) \} = m \text{Tr} \mathbf{H} + \frac{1}{2} \left(\text{Tr}(\mathbf{Z}' \nabla_Z (\mathbf{U} \Phi \mathbf{U}^*)) + \text{Tr}(\mathbf{Z}^* \nabla_{\bar{Z}} \mathbf{U} \Phi \mathbf{U}^*) \right), \quad (15)$$

where $\nabla_{\bar{Z}} = (\partial/\partial \bar{z}_{jk})_{j=1, 2, \dots, m, k=1, 2, \dots, p}$ with $\partial/\partial \bar{z}_{jk} = \{\partial/\partial(\text{Re } z_{jk}) + \sqrt{-1} \partial/\partial(\text{Im } z_{jk})\}/2$. We

use Lemma 5 to evaluate the second term in the right hand side of (15) as

$$\begin{aligned}
\text{Tr}(\mathbf{Z}'\nabla_Z(\mathbf{U}\Phi\mathbf{U}^*))' &= \sum_{i,j,k,l} z_{ji} \frac{\partial(\bar{u}_{kl}\varphi_l u_{il})}{\partial z_{jk}} \\
&= \sum_{i,j,k,l} z_{ji} \left\{ \varphi_l u_{il} \frac{\partial \bar{u}_{kl}}{\partial z_{jk}} + \varphi_l \bar{u}_{kl} \frac{\partial u_{il}}{\partial z_{jk}} + \bar{u}_{kl} u_{il} \sum_{k'} \frac{\partial \varphi_l}{\partial \ell_{k'}} \frac{\partial \ell_{k'}}{\partial z_{jk}} \right\} \\
&= \sum_{i,j,k,l} z_{ji} \left\{ \varphi_l u_{il} \sum_b \sum_{c \neq l} \frac{\bar{u}_{kc} u_{kc} \bar{u}_{bl} \bar{z}_{jb}}{\ell_l - \ell_c} + \varphi_l \bar{u}_{kl} \sum_b \sum_{c \neq l} \frac{u_{ic} \bar{u}_{bc} u_{kl} \bar{z}_{jb}}{\ell_l - \ell_c} \right. \\
&\quad \left. + \bar{u}_{kl} u_{il} \sum_{k',b} u_{kk'} \bar{u}_{bk'} \bar{z}_{jb} \frac{\partial \varphi_l}{\partial \ell_{k'}} \right\} \\
&= \sum_k \left\{ \sum_{c \neq k} \frac{\ell_k \varphi_k - \ell_c \varphi_c}{\ell_k - \ell_c} + \ell_k \frac{\partial \varphi_k}{\partial \ell_k} \right\}.
\end{aligned}$$

We use Lemma 5 and the fact that $\partial q / \partial \bar{z} = \overline{\partial \bar{q} / \partial z}$ to evaluate the third term in the right hand side of (15) as

$$\begin{aligned}
\text{Tr}(\mathbf{Z}^* \nabla_{\bar{Z}} \mathbf{U} \Phi \mathbf{U}^*) &= \sum_{i,j,k,l} \bar{z}_{ji} \frac{\partial(u_{kl}\varphi_l \bar{u}_{il})}{\partial \bar{z}_{jk}} \\
&= \sum_{i,j,k,l} \bar{z}_{ji} \left\{ \varphi_l \bar{u}_{il} \frac{\partial \bar{u}_{kl}}{\partial \bar{z}_{jk}} + \varphi_l u_{kl} \frac{\partial \bar{u}_{il}}{\partial \bar{z}_{jk}} + u_{kl} \bar{u}_{il} \sum_{k'} \frac{\partial \varphi_l}{\partial \ell_{k'}} \frac{\partial \ell_{k'}}{\partial \bar{z}_{jk}} \right\} \\
&= \sum_k \left\{ \sum_{c \neq k} \frac{\ell_k \varphi_k - \ell_c \varphi_c}{\ell_k - \ell_c} + \ell_k \frac{\partial \varphi_k}{\partial \ell_k} \right\}.
\end{aligned}$$

Putting the above two equations into (15), we have

$$\begin{aligned}
\text{Tr}\{\text{Re}(\nabla_Z' \mathbf{Z} \mathbf{U} \Phi \mathbf{U}^*)\} &= \sum_k \left\{ m \varphi_k + \sum_{c \neq k} \frac{\ell_k \varphi_k - \ell_c \varphi_c + \varphi_c (\ell_c - \ell_k)}{\ell_k - \ell_c} + \ell_k \frac{\partial \varphi_k}{\partial \ell_k} \right\} \\
&= \sum_k \left\{ (m - p + 1) \varphi_k + \sum_{c \neq k} \frac{\ell_k \varphi_k - \ell_c \varphi_c}{\ell_k - \ell_c} + \ell_k \frac{\partial \varphi_k}{\partial \ell_k} \right\}.
\end{aligned}$$

Combining this equation with (15), we completes the proof of this lemma. \square

A.2 Eigencalculus for unknown covariance case with $m > p$

Next we record calculus on the eigenvalues for the case when $m > p$. Let $\mathbf{A} = (a_{ij})_{i,j=1,2,\dots,p}$ be a $p \times p$ nonsingular matrix such that

$$\mathbf{A}^* \mathbf{S} \mathbf{A} = \mathbf{I}_p, \quad \mathbf{A}^* \mathbf{Z}^* \mathbf{Z} \mathbf{A} = \mathbf{F}, \quad \mathbf{F} = \text{Diag}(f_1, f_2, \dots, f_p)$$

with $f_1 > f_2 > \dots > f_p > 0$. This means that we consider (\mathbf{Z}, \mathbf{S}) such that a matrix $\mathbf{Z}^* \mathbf{Z} \mathbf{S}^{-1}$ has the distinct eigenvalues f_1, f_2, \dots, f_p .

Lemma 7. *Let $\mathbf{A}^{-1} = (a^{ij})_{i,j=1,2,\dots,p}$, $\bar{\mathbf{A}} = (\bar{a}_{ij})_{i,j=1,2,\dots,p}$, and $(\bar{\mathbf{A}})^{-1} = (\bar{a}^{ij})_{i,j=1,2,\dots,p}$. For $i, k, k' = 1, 2, \dots, p$, and $j = 1, 2, \dots, m$, we have*

$$\begin{aligned} \frac{\partial a^{lk'}}{\partial z_{jk}} &= \sum_b \sum_{c \neq l} \frac{\bar{a}_{bl} a_{kc} a^{ck'} \bar{z}_{jb}}{f_l - f_c}, \\ \frac{\partial a_{il}}{\partial z_{jk}} &= \sum_b \sum_{c \neq l} \frac{a_{ic} \bar{a}_{bc} a_{kl} \bar{z}_{jb}}{f_l - f_c}, \\ \frac{\partial f_{k'}}{\partial z_{jk}} &= \sum_b \bar{a}_{bk'} a_{kk'} \bar{z}_{jb}, \\ \frac{\partial (a^{ki} \bar{a}^{kj})}{\partial s_{ij}} &= a^{ki} \bar{a}^{kj} \bar{a}_{jk} a_{ik} + \sum_{b \neq k} a^{ki} a_{ik} \bar{a}_{jb} \bar{a}^{bj} \frac{f_b}{f_b - f_k} + \sum_{b \neq k} \bar{a}^{kj} \bar{a}_{jk} a_{ib} a^{bi} \frac{f_b}{f_b - f_k}, \\ \frac{\partial f_i}{\partial s_{kk'}} &= -\bar{a}_{k'i} a_{ki} f_i. \end{aligned}$$

Proof. Put $\mathbf{W} = \mathbf{Z}^* \mathbf{Z} = (w_{ij})_{i,j=1,2,\dots,p}$. Differentiating $\mathbf{S} = (\mathbf{A}^*)^{-1} \mathbf{A}^{-1}$ and $\mathbf{W} = (\mathbf{A}^*)^{-1} \mathbf{F} \mathbf{A}^{-1}$, we have

$$\begin{aligned} (d\mathbf{S}) &= (\mathbf{A}^*)^{-1} (d\mathbf{A}^{-1}) + (d(\mathbf{A}^*)^{-1}) \mathbf{A}^{-1} \\ (d\mathbf{W}) &= (\mathbf{A}^*)^{-1} \mathbf{F} (d\mathbf{A}^{-1}) + (d(\mathbf{A}^*)^{-1}) \mathbf{F} \mathbf{A}^{-1} + (\mathbf{A}^*)^{-1} (d\mathbf{F}) \mathbf{A}^{-1}. \end{aligned}$$

Multiplying these equations by \mathbf{A}^* on the left and by \mathbf{A} on the right, we get

$$\mathbf{A}^* (d\mathbf{S}) \mathbf{A} = (d\mathbf{A}^{-1}) \mathbf{A} + \mathbf{A}^* (d(\mathbf{A}^*)^{-1}) \quad (16)$$

$$\mathbf{A}^* (d\mathbf{W}) \mathbf{A} = \mathbf{F} (d\mathbf{A}^{-1}) \mathbf{A} + \mathbf{A}^* (d(\mathbf{A}^*)^{-1}) \mathbf{F} + (d\mathbf{F}). \quad (17)$$

To obtain the derivatives with respect to $\partial/\partial z_{jk}$, we may assume that $d\mathbf{S} = 0$. Then, putting (16) into (17) and from some algebraic calculation, we have

$$(d\mathbf{F})_{k'k'} = \sum_{b,c} \bar{a}_{bk'} a_{ck'} (d\mathbf{W})_{bc} \quad (18)$$

and

$$((d\mathbf{A}^{-1}) \mathbf{A})_{lb} = \begin{cases} \frac{1}{f_l - f_b} \sum_{c_1, c_2} \bar{a}_{c_1 l} a_{c_2 b} (d\mathbf{W})_{c_1 c_2} & \text{if } l \neq b, \\ 0 & \text{if } l = b. \end{cases} \quad (19)$$

From (19) and the fact that

$$\begin{aligned} (\mathrm{d}\mathbf{W})_{c_1 c_2} \left(\frac{\partial}{\partial z_{jk}} \right) &= \sum_{c_3} \left\{ z_{c_3 c_2} (\mathrm{d}\bar{\mathbf{Z}})_{c_3 c_1} \left(\frac{\partial}{\partial z_{jk}} \right) + \bar{z}_{c_3 c_1} (\mathrm{d}\mathbf{Z})_{c_3 c_2} \left(\frac{\partial}{\partial z_{jk}} \right) \right\} \\ &= \sum_{c_3} \bar{z}_{c_3 c_1} \delta_{c_3 j} \delta_{c_2 k} = \bar{z}_{j c_1} \delta_{c_2 k}, \end{aligned}$$

we have

$$\begin{aligned} \frac{\partial a^{lk'}}{\partial z_{jk}} &= (\mathrm{d}\mathbf{A}^{-1})_{lk'} \left(\frac{\partial}{\partial z_{jk}} \right) = \sum_b a^{bk'} ((\mathrm{d}\mathbf{A}^{-1})\mathbf{A})_{lb} \left(\frac{\partial}{\partial z_{jk}} \right) \\ &= \sum_{c_1, c_2} \sum_{b \neq l} \frac{\bar{a}_{c_1 l} a_{c_2 b} a^{bk'}}{f_l - f_b} (\mathrm{d}\mathbf{W})_{c_1 c_2} \left(\frac{\partial}{\partial z_{jk}} \right) \\ &= \sum_{c_1} \sum_{b \neq l} \frac{\bar{a}_{c_1 l} a_{kb} a^{bk'} \bar{z}_{j c_1}}{f_l - f_b}, \end{aligned}$$

which completes the first equation of this lemma.

Differentiating $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_p$, we have $(\mathrm{d}\mathbf{A})\mathbf{A}^{-1} + \mathbf{A}(\mathrm{d}\mathbf{A}^{-1}) = \mathbf{0}$. Multiplying this equation by \mathbf{A} on the right and using the first assertion of this lemma, we have

$$\begin{aligned} \frac{\partial a_{il}}{\partial z_{jk}} &= - \sum_{b_1, b_2} a_{ib_1} a_{b_2 l} \frac{\partial a^{b_1 b_2}}{\partial z_{jk}} = - \sum_{b_1, b_2} a_{ib_1} a_{b_2 l} \sum_{b_4} \sum_{b_3 \neq b_1} \frac{\bar{a}_{b_4 b_1} a_{kb_3} a^{b_3 b_2} \bar{z}_{j b_4}}{f_{b_1} - f_{b_3}} \\ &= - \sum_{b_4} \sum_{b_1 \neq l} \frac{a_{ib_1} \bar{a}_{b_4 b_1} a_{kl} \bar{z}_{j b_4}}{f_{b_1} - f_l}, \end{aligned}$$

which completes the proof of the second assertion.

From (18) we have

$$\frac{\partial f_{k'}}{\partial z_{jk}} = \sum_{b_1, b_2} \bar{a}_{b_1 k'} a_{b_2 k'} (\mathrm{d}\mathbf{W})_{b_1 b_2} \left(\frac{\partial}{\partial z_{jk}} \right) = \sum_{b_1, b_2} \bar{a}_{b_1 k'} a_{b_2 k'} \bar{z}_{j b_1} \delta_{b_2 k} = \sum_{b_1} \bar{a}_{b_1 k'} a_{k k'} \bar{z}_{j b_1},$$

which completes the third assertion of this lemma.

To derive the derivatives with respect to s_{ij} we assume that $\mathrm{d}\mathbf{W} = 0$ in (16). Reverting to the coordinates, we have

$$(\mathrm{d}\mathbf{F})_{ii} = - \sum_{j, k} \bar{a}_{ji} a_{ki} f_i (\mathrm{d}\mathbf{S})_{jk} \quad (20)$$

and

$$((\mathrm{d}\mathbf{A}^{-1})\mathbf{A})_{ij} = \frac{f_j}{f_j - f_i} (\mathbf{A}^* (\mathrm{d}\mathbf{S}) \mathbf{A})_{ij}, \quad \text{if } i \neq j, \quad (21)$$

$$(\overline{(\mathrm{d}\mathbf{A}^{-1})\mathbf{A}})_{ij} = \frac{f_j}{f_j - f_i} (\overline{\mathbf{A}^* (\mathrm{d}\mathbf{S}) \mathbf{A}})_{ij}, \quad \text{if } i \neq j. \quad (22)$$

Since $\mathbf{A}^*(d\mathbf{S})\mathbf{A} = (d\mathbf{A}^{-1})\mathbf{A} + \mathbf{A}^*(d(\mathbf{A}^*)^{-1})$ which implies that $2\text{Re} [((d\mathbf{A}^{-1})\mathbf{A})_{ii}] = (\mathbf{A}^*(d\mathbf{S})\mathbf{A})_{ii}$, we have

$$((d\mathbf{A}^{-1})\mathbf{A})_{ii} + (\overline{((d\mathbf{A}^{-1})\mathbf{A})})_{ii} = (\mathbf{A}^*(d\mathbf{S})\mathbf{A})_{ii}. \quad (23)$$

From (21) – (23) we have

$$\begin{aligned} \frac{\partial(a^{ki}\bar{a}^{kj})}{\partial s_{ij}} &= a^{ki}\frac{\partial\bar{a}^{kj}}{\partial s_{ij}} + \bar{a}^{kj}\frac{\partial a^{ki}}{\partial s_{ij}} = a^{ki}(\overline{(d\mathbf{A}^{-1})})_{kj} \left(\frac{\partial}{\partial s_{ij}}\right) + \bar{a}^{kj}(d\mathbf{A}^{-1})_{ki} \left(\frac{\partial}{\partial s_{ij}}\right) \\ &= a^{ki} \sum_{b_1} \bar{a}^{b_1j} (\overline{(d\mathbf{A}^{-1})\mathbf{A}})_{kb_1} \left(\frac{\partial}{\partial s_{ij}}\right) + \bar{a}^{kj} \sum_{b_1} a^{b_1i} ((d\mathbf{A}^{-1})\mathbf{A})_{kb_1} \left(\frac{\partial}{\partial s_{ij}}\right) \\ &= a^{ki}\bar{a}^{kj} \left\{ (\overline{(d\mathbf{A}^{-1})\mathbf{A}})_{kk} \left(\frac{\partial}{\partial s_{ij}}\right) + ((d\mathbf{A}^{-1})\mathbf{A})_{kk} \left(\frac{\partial}{\partial s_{ij}}\right) \right\} \\ &\quad + a^{ki} \sum_{b_1 \neq k} \bar{a}^{b_1j} (\overline{(d\mathbf{A}^{-1})\mathbf{A}})_{kb_1} \left(\frac{\partial}{\partial s_{ij}}\right) + \bar{a}^{kj} \sum_{b_1 \neq k} a^{b_1i} ((d\mathbf{A}^{-1})\mathbf{A})_{kb_1} \left(\frac{\partial}{\partial s_{ij}}\right) \\ &= a^{ki}\bar{a}^{kj} \sum_{b_2, b_3} \bar{a}_{b_2k} a_{b_3k} (d\mathbf{S})_{b_2b_3} \left(\frac{\partial}{\partial s_{ij}}\right) \\ &\quad + a^{ki} \sum_{b_2, b_3} \sum_{b_1 \neq k} a_{b_2k} \bar{a}_{b_3b_1} \bar{a}^{b_1j} \frac{f_{b_1}}{f_{b_1} - f_k} (d\bar{\mathbf{S}})_{b_2b_3} \left(\frac{\partial}{\partial s_{ij}}\right) \\ &\quad + \bar{a}^{kj} \sum_{b_2, b_3} \sum_{b_1 \neq k} \bar{a}_{b_2k} a_{b_3b_1} a^{b_1i} \frac{f_{b_1}}{f_{b_1} - f_k} (d\mathbf{S})_{b_2b_3} \left(\frac{\partial}{\partial s_{ij}}\right) \\ &= a^{ki}\bar{a}^{kj} \sum_{b_2, b_3} \bar{a}_{b_2k} a_{b_3k} \delta_{b_2j} \delta_{b_3i} + a^{ki} \sum_{b_2, b_3} \sum_{b_1 \neq k} a_{b_2k} \bar{a}_{b_3b_1} \bar{a}^{b_1j} \frac{f_{b_1}}{f_{b_1} - f_k} \delta_{b_2i} \delta_{b_3j} \\ &\quad + \bar{a}^{kj} \sum_{b_2, b_3} \sum_{b_1 \neq k} \bar{a}_{b_2k} a_{b_3b_1} a^{b_1i} \frac{f_{b_1}}{f_{b_1} - f_k} \delta_{b_2j} \delta_{b_3i} \\ &= a^{ki}\bar{a}^{kj} \bar{a}_{jk} a_{ik} + \sum_{b_1 \neq k} a^{ki} a_{ik} \bar{a}_{jb_1} \bar{a}^{b_1j} \frac{f_{b_1}}{f_{b_1} - f_k} \\ &\quad + \sum_{b_1 \neq k} \bar{a}^{kj} \bar{a}_{jk} a_{ib_1} a^{b_1i} \frac{f_{b_1}}{f_{b_1} - f_k}, \end{aligned}$$

which completes the proof of forth assertion of this lemma.

Finally, from (20), we have

$$\begin{aligned} \frac{\partial f_{b_1}}{\partial s_{ij}} &= (d\mathbf{F})_{b_1b_1} \left(\frac{\partial}{\partial s_{ij}}\right) = - \sum_{b_2, b_3} \bar{a}_{b_2b_1} a_{b_3b_1} f_{b_1} (d\mathbf{S})_{b_2b_3} \left(\frac{\partial}{\partial s_{ij}}\right) = - \sum_{b_2, b_3} \bar{a}_{b_2b_1} a_{b_3b_1} f_{b_1} \delta_{b_2j} \delta_{b_3i} \\ &= -\bar{a}_{jb_1} a_{ib_1} f_{b_1}, \end{aligned}$$

which completes the final part of this lemma. \square

Lemma 8. Let $\Phi(\mathbf{F}) = \text{Diag}(\varphi_1(\mathbf{F}), \varphi_2(\mathbf{F}), \dots, \varphi_p(\mathbf{F}))$, where $\varphi_i(\mathbf{F})$'s ($i = 1, 2, \dots, p$) are differentiable functions from $\mathbb{R}_{>}^p \rightarrow \mathbb{R}_+$. Then we have

$$\begin{aligned} \text{Tr}\{\text{Re}(\nabla'_Z \mathbf{Z} \mathbf{A} \Phi \mathbf{A}^{-1})\} &= \sum_k \left\{ f_k \varphi_{kk}(\mathbf{F}) + (m - p + 1) \varphi_k(\mathbf{F}) + 2 \sum_{b>k} \frac{f_k \varphi_k(\mathbf{F}) - f_b \varphi_b(\mathbf{F})}{f_k - f_b} \right\}, \\ \text{Tr}(\mathbf{D}_S((\mathbf{A}^*)^{-1} \Phi(\mathbf{F}) \mathbf{A}^{-1})) &= \sum_k \left\{ (2p - 1) \varphi_k(\mathbf{F}) - 2 \sum_{b>k} \frac{f_k \varphi_k(\mathbf{F}) - f_b \varphi_b(\mathbf{F})}{f_k - f_b} - f_k \varphi_{kk}(\mathbf{F}) \right\}, \end{aligned}$$

where $\varphi_{kk}(\mathbf{F}) = (\partial \varphi_k / \partial f_k)(\mathbf{F})$, $k = 1, 2, \dots, p$.

Proof. Use notation Φ , φ_k , and φ_{kk} short for $\Phi(\mathbf{F})$, $\varphi_k(\mathbf{F})$, and $\varphi_{kk}(\mathbf{F})$, respectively. To prove the first equation of this lemma, we first note that

$$\text{Tr}\{\text{Re}(\nabla'_Z \mathbf{Z} \mathbf{A} \Phi \mathbf{A}^{-1})\} = m \text{Tr} \Phi + \frac{1}{2} \left(\text{Tr}\{\mathbf{Z}' \nabla_Z ((\mathbf{A}')^{-1} \Phi \mathbf{A}')\} + \text{Tr}\{\mathbf{Z}^* \nabla_{\bar{\mathbf{Z}}} ((\mathbf{A}^*)^{-1} \Phi \mathbf{A}^*)\} \right).$$

Now we use the first three equations in Lemma 7 to evaluate the second term inside expectation of the right hand side as

$$\begin{aligned} \text{Tr}\{\mathbf{Z}' \nabla_Z ((\mathbf{A}')^{-1} \Phi \mathbf{A}')\} &= \sum_{i,j,k,l} z_{ji} \frac{\partial(a^{lk} \varphi_l a_{il})}{\partial z_{jk}} \\ &= \sum_{i,j,k,l} z_{ji} \left\{ \varphi_l a_{il} \frac{\partial a^{lk}}{\partial z_{jk}} + \varphi_l a^{lk} \frac{\partial a_{il}}{\partial z_{jk}} + a^{lk} a_{il} \frac{\partial \varphi_l}{\partial z_{jk}} \right\} \\ &= \sum_{i,j,k,l} \left\{ z_{ji} \varphi_l a_{il} \sum_{b_2} \sum_{b_1 \neq l} \frac{\bar{a}_{b_2 l} a_{k b_1} a^{b_1 k} \bar{z}_{j b_2}}{f_l - f_{b_1}} + z_{ji} \varphi_l a^{lk} \sum_{b_2} \sum_{b_1 \neq l} \frac{a_{i b_1} \bar{a}_{b_2 b_1} a_{k l} \bar{z}_{j b_2}}{f_l - f_{b_1}} \right. \\ &\quad \left. + z_{ji} a^{lk} a_{il} \sum_{b_1, b_2} \bar{a}_{b_2 b_1} a_{k b_1} \bar{z}_{j b_2} \frac{\partial \varphi_l}{\partial f_{b_1}} \right\} \\ &= \sum_{i,j,l,b_2} \left\{ \sum_{b_1 \neq l} \bar{z}_{j b_2} z_{ji} a_{il} \bar{a}_{b_2 l} \frac{\varphi_l}{f_l - f_{b_1}} + \sum_{b_1 \neq l} \bar{z}_{j b_2} z_{ji} a_{i b_1} \bar{a}_{b_2 b_1} \frac{\varphi_l}{f_l - f_{b_1}} \right. \\ &\quad \left. + \bar{z}_{j b_2} z_{ji} a_{il} \bar{a}_{b_2 l} \frac{\partial \varphi_l}{\partial f_l} \right\} \\ &= \text{Tr}(\mathbf{Z}^* \mathbf{Z} \mathbf{A} \tilde{\Phi} \mathbf{A}^*), \end{aligned}$$

where $\tilde{\Phi} = \text{Diag}(\tilde{\phi}_1, \tilde{\phi}_2, \dots, \tilde{\phi}_p)$ with $\tilde{\phi}_i = (\partial \varphi_i / \partial f_i) + \sum_{b \neq i} (\varphi_i - \varphi_b) / (f_i - f_b)$, ($i = 1, 2, \dots, p$).

Similarly we use the first three equations in Lemma 7 to evaluate the third term inside expectation

tation of the right hand side as

$$\begin{aligned}
\text{Tr} \{ \mathbf{Z}^* \nabla_{\bar{\mathbf{Z}}} ((\mathbf{A}^*)^{-1} \Phi \mathbf{A}^*) \} &= \sum_{i,j,k,l} \bar{z}_{ji} \frac{\partial(\bar{a}^{lk} \varphi_l \bar{a}_{il})}{\partial \bar{z}_{jk}} \\
&= \sum_{i,j,k,l} \bar{z}_{ji} \left\{ \varphi_l \bar{a}_{il} \frac{\partial \bar{a}^{lk}}{\partial z_{jk}} + \varphi_l \bar{a}^{lk} \frac{\partial \bar{a}_{il}}{\partial z_{jk}} + \bar{a}^{lk} \bar{a}_{il} \frac{\partial \varphi_l}{\partial z_{jk}} \right\} \\
&= \text{Tr} (\mathbf{Z}^* \mathbf{Z} \mathbf{A} \tilde{\Phi} \mathbf{A}^*).
\end{aligned}$$

Putting $\mathbf{Z}^* \mathbf{Z} = (\mathbf{A}^*)^{-1} \mathbf{F} \mathbf{A}$ into the right hand side of the above two equations, we have

$$\text{Tr} \{ \text{Re} (\nabla'_{\mathbf{Z}} \mathbf{Z} \mathbf{A} \Phi \mathbf{A}^{-1}) \} = \sum_k \left\{ f_k \varphi_{kk} + (m - p + 1) \varphi_k + \sum_{b \neq k} \frac{f_k \varphi_k - f_b \varphi_b}{f_k - f_b} \right\},$$

which completes the first equation of this lemma.

Next we prove the second equation of this lemma. Apply the chain rule first and use the forth and fifth equations of Lemma 7 to get

$$\begin{aligned}
\text{Tr} (\mathbf{D}_S ((\mathbf{A}^*)^{-1} \Phi \mathbf{A}^{-1})) &= \sum_{i,j,k} \frac{\partial(\bar{a}^{kj} \varphi_k a^{ki})}{\partial s_{ij}} \\
&= \sum_k \left\{ \varphi_k \sum_{i,j} \frac{\partial(\bar{a}^{kj} a^{ki})}{\partial s_{ij}} + \sum_{i,j} a^{ki} \bar{a}^{kj} \sum_b \frac{\partial \varphi_k}{\partial f_b} \frac{\partial f_b}{\partial s_{ij}} \right\} \\
&= \sum_k \left\{ \varphi_k \left(1 + 2 \sum_{b \neq k} \frac{f_b}{f_b - f_k} \right) - f_k \frac{\partial \varphi_k}{\partial f_k} \right\} \\
&= \sum_k \left\{ \varphi_k + 2 \sum_{b \neq k} \frac{(f_b - f_k) \varphi_k + f_k \varphi_k}{f_b - f_k} - f_k \varphi_{kk} \right\} \\
&= \sum_k \left\{ \varphi_k + 2(p - 1) \varphi_k - 2 \sum_{b > k} \frac{f_k \varphi_k - f_b \varphi_b}{f_k - f_b} - f_k \varphi_{kk} \right\},
\end{aligned}$$

which completes the proof of this lemma. \square

A.3 Eigencalculus for unknown covariance case with $m < p$

Let $\mathbf{Z}^* \mathbf{S}^{-1} \mathbf{Z} = \mathbf{U} \mathbf{F} \mathbf{U}^*$, where $\mathbf{U} = (u_{ij})_{i,j=1,\dots,m}$ is an $m \times m$ unitary matrix and $\mathbf{F} = \text{Diag}(f_1, f_2, \dots, f_m)$ with diagonal elements f_1, f_2, \dots, f_m ($f_1 > f_2 > \dots > f_m > 0$).

Lemma 9. For $i, k, l, l' = 1, 2, \dots, p, j = 1, 2, \dots, m$, we have

$$\begin{aligned}
\frac{\partial u_{il}}{\partial z_{jk}} &= \sum_{b_3, b_5} \sum_{b_1 \neq l} \frac{u_{ib_1} \bar{u}_{jb_1} u_{b_3 l} s^{kb_5} \bar{z}_{b_3 b_5}}{f_l - f_{b_1}}, \\
\frac{\partial \bar{u}_{il}}{\partial z_{jk}} &= \sum_{b_3, b_5} \sum_{b_1 \neq l} \frac{\bar{u}_{ib_1} \bar{u}_{jb_1} u_{b_3 b_1} s^{kb_5} \bar{z}_{b_3 b_5}}{f_l - f_{b_1}}, \\
\frac{\partial f_{b_1}}{\partial z_{jk}} &= \sum_{b_3, b_5} \bar{u}_{jb_1} u_{b_3 b_1} s^{kb_5} \bar{z}_{b_3 b_5}, \\
\frac{\partial u_{kl}}{\partial s_{ij}} &= - \sum_{b_2, b_3, b_4, b_5} \sum_{b_1 \neq l} \frac{u_{kb_1} \bar{u}_{b_2 b_1} u_{b_3 l} z_{b_2 b_4} \bar{z}_{b_3 b_5} s^{b_4 j} s^{ib_5}}{f_l - b_{b_1}}, \\
\frac{\partial \bar{u}_{kl}}{\partial s_{ij}} &= - \sum_{b_2, b_3, b_4, b_5} \sum_{b_1 \neq l} \frac{\bar{u}_{kb_1} \bar{u}_{b_2 l} u_{b_3 b_1} z_{b_2 b_4} \bar{z}_{b_3 b_5} s^{b_4 j} s^{ib_5}}{f_l - b_{b_1}}, \\
\frac{\partial f_{l'}}{\partial s_{ij}} &= - \sum_{b_2, b_3, b_4, b_5} \bar{u}_{b_2 l'} u_{b_3 l'} z_{b_2 b_4} s^{b_4 j} s^{ib_5} \bar{z}_{b_3 b_5}.
\end{aligned}$$

Proof. Let $\mathbf{T} = \mathbf{ZS}^{-1}\mathbf{Z}^*$ and take differential of $\mathbf{T} = \mathbf{U}\mathbf{F}\mathbf{U}^*$ to get

$$(\mathrm{d}\mathbf{T}) = (\mathrm{d}\mathbf{U})\mathbf{F}\mathbf{U}^* + \mathbf{U}\mathbf{F}(\mathrm{d}\mathbf{U}^*) + \mathbf{U}(\mathrm{d}\mathbf{F})\mathbf{U}^*,$$

from which it follows that

$$\mathbf{U}^*(\mathrm{d}\mathbf{T})\mathbf{U} = \mathbf{U}^*(\mathrm{d}\mathbf{U})\mathbf{F} + \mathbf{F}(\mathrm{d}\mathbf{U}^*)\mathbf{U} + (\mathrm{d}\mathbf{F}) = (\mathbf{U}^*(\mathrm{d}\mathbf{U}))\mathbf{F} - \mathbf{F}(\mathbf{U}^*(\mathrm{d}\mathbf{U})) + (\mathrm{d}\mathbf{F}).$$

Therefore we have

$$(\mathbf{U}^*(\mathrm{d}\mathbf{U}))_{ij} = \begin{cases} \frac{1}{f_j - f_i} (\mathbf{U}^*(\mathrm{d}\mathbf{T})\mathbf{U})_{ij} & \text{if } i \neq j \\ 0 & \text{if } i = j, \end{cases} \quad (24)$$

while, for $i = j$, we have

$$(\mathrm{d}\mathbf{F})_{ii} = (\mathbf{U}^*(\mathrm{d}\mathbf{T})\mathbf{U})_{ii}. \quad (25)$$

Hence

$$\begin{aligned}
\frac{\partial u_{il}}{\partial z_{jk}} &= (\mathrm{d}\mathbf{U})_{il} \left(\frac{\partial}{\partial z_{jk}} \right) = \sum_{b_1} u_{ib_1} (\mathbf{U}^*(\mathrm{d}\mathbf{U}))_{b_1 l} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{b_2, b_3} \sum_{b_1 \neq l} \frac{u_{ib_1} \bar{u}_{b_2 b_1} u_{b_3 l}}{f_l - f_{b_1}} (\mathrm{d}\mathbf{T})_{b_2 b_3} \left(\frac{\partial}{\partial z_{jk}} \right).
\end{aligned}$$

But

$$\begin{aligned}
(\mathbf{d}\mathbf{T})_{b_2b_3} \left(\frac{\partial}{\partial z_{jk}} \right) &= \sum_{b_4, b_5} \left\{ \bar{z}_{b_3b_5} s^{b_4b_5} (\mathbf{d}\mathbf{Z})_{b_2b_4} \left(\frac{\partial}{\partial z_{jk}} \right) + z_{b_2b_4} s^{b_4b_5} (\mathbf{d}\bar{\mathbf{Z}})_{b_3b_5} \left(\frac{\partial}{\partial z_{jk}} \right) \right\} \\
&= \sum_{b_4, b_5} \delta_{jb_2} \delta_{kb_4} s^{b_4b_5} \bar{z}_{b_3b_5} = \sum_{b_5} \delta_{jb_2} s^{kb_5} \bar{z}_{b_3b_5}.
\end{aligned} \tag{26}$$

Therefore we have

$$\frac{\partial u_{il}}{\partial z_{jk}} = \sum_{b_2, b_3, b_5} \sum_{b_1 \neq l} \frac{u_{ib_1} \bar{u}_{b_2b_1} u_{b_3l} \delta_{jb_2} s^{kb_5} \bar{z}_{b_3b_5}}{f_l - f_{b_1}} = \sum_{b_3, b_5} \sum_{b_1 \neq l} \frac{u_{ib_1} \bar{u}_{jb_1} u_{b_3l} s^{kb_5} \bar{z}_{b_3b_5}}{f_l - f_{b_1}},$$

which complete the proof of the first assertion of this lemma.

Similarly we have $((\mathbf{d}\mathbf{U}^*)\mathbf{U})_{ij} = (\mathbf{U}^*(\mathbf{d}\mathbf{T})\mathbf{U})_{ij}/(f_i - f_j)$ for $i \neq j$, from which it follows that

$$\begin{aligned}
\frac{\partial \bar{u}_{il}}{\partial z_{jk}} &= (\mathbf{d}\mathbf{U}^*)_{li} \left(\frac{\partial}{\partial z_{jk}} \right) = \sum_{b_1} \bar{u}_{ib_1} ((\mathbf{d}\mathbf{U}^*)\mathbf{U})_{lb_1} \left(\frac{\partial}{\partial z_{jk}} \right) = \sum_{b_2, b_3} \sum_{b_1 \neq l} \frac{\bar{u}_{ib_1} \bar{u}_{b_2l} u_{b_3b_1}}{f_l - f_{b_1}} (\mathbf{d}\mathbf{T})_{b_2b_3} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{b_2, b_3, b_5} \sum_{b_1 \neq l} \frac{\bar{u}_{ib_1} \bar{u}_{b_2l} u_{b_3b_1} \delta_{jb_2} s^{kb_5} \bar{z}_{b_3b_5}}{f_l - f_{b_1}} \\
&= \sum_{b_3, b_5} \sum_{b_1 \neq l} \frac{\bar{u}_{ib_1} \bar{u}_{jl} u_{b_3b_1} s^{kb_5} \bar{z}_{b_3b_5}}{f_l - f_{b_1}},
\end{aligned}$$

which completes the second assertion of this lemma.

Furthermore, from (25) and (26), we have

$$\begin{aligned}
\frac{\partial f_{b_1}}{\partial z_{jk}} &= (\mathbf{d}\mathbf{F})_{b_1b_1} \left(\frac{\partial}{\partial z_{jk}} \right) = (\mathbf{U}^*(\mathbf{d}\mathbf{T})\mathbf{U})_{b_1b_1} \left(\frac{\partial}{\partial z_{jk}} \right) = \sum_{b_2, b_3} \bar{u}_{b_2b_1} u_{b_3b_1} (\mathbf{d}\mathbf{T})_{b_2b_3} \left(\frac{\partial}{\partial z_{jk}} \right) \\
&= \sum_{b_2, b_3, b_5} \bar{u}_{b_2b_1} u_{b_3b_1} \delta_{jb_2} s^{kb_5} \bar{z}_{b_3b_5} = \sum_{b_3, b_5} \bar{u}_{jb_1} u_{b_3b_1} s^{kb_5} \bar{z}_{b_3b_5},
\end{aligned}$$

which complete the proof of the third assertion of this lemma.

Next we prove the forth equation of this lemma. Use (24) to get

$$\begin{aligned}
\frac{\partial u_{kl}}{\partial s_{ij}} &= \sum_{b_1 \neq l} u_{kb_1} (\mathbf{U}^*(\mathbf{d}\mathbf{U}))_{b_1l} \left(\frac{\partial}{\partial s_{ij}} \right) \\
&= \sum_{b_2, b_3} \sum_{b_1 \neq l} \frac{u_{kb_1} \bar{u}_{b_2b_1} u_{b_3l}}{f_l - f_{b_1}} (\mathbf{d}\mathbf{T})_{b_2b_3} \left(\frac{\partial}{\partial s_{ij}} \right).
\end{aligned}$$

But, since $(\partial s^{b_4b_5}/\partial s_{ij}) = -s^{b_4j} s^{ib_5}$, we have

$$(\mathbf{d}\mathbf{T})_{b_2b_3} \left(\frac{\partial}{\partial s_{ij}} \right) = \sum_{b_4, b_5} z_{b_2b_4} \bar{z}_{b_3b_5} (\mathbf{d}\mathbf{S}^{-1})_{b_4b_5} \left(\frac{\partial}{\partial s_{ij}} \right) = - \sum_{b_4, b_5} z_{b_2b_4} s^{b_4j} s^{ib_5} \bar{z}_{b_3b_5}.$$

Combining these two equations we completes the proof of the forth equation of this lemma.

Similarly we have

$$\begin{aligned}
\frac{\partial \bar{u}_{kl}}{\partial s_{ij}} &= \sum_{b_1} \bar{u}_{kb_1} ((d\mathbf{U}^*)\mathbf{U})_{lb_1} \left(\frac{\partial}{\partial s_{ij}} \right) \\
&= \sum_{b_2, b_3} \sum_{b_1 \neq l} \frac{\bar{u}_{kb_1} \bar{u}_{b_2 l} u_{b_3 b_1}}{f_l - f_{b_1}} (d\mathbf{T})_{b_2 b_3} \left(\frac{\partial}{\partial s_{ij}} \right) \\
&= - \sum_{b_2, b_3, b_4, b_5} \sum_{b_1 \neq l} \frac{\bar{u}_{kb_1} \bar{u}_{b_2 l} u_{b_3 b_1} z_{b_2 b_4} \bar{z}_{b_3 b_5} s^{b_4 j} s^{ib_5}}{f_l - f_{b_1}},
\end{aligned}$$

which completes the proof of the forth assertion of this lemma.

Finally, from (25), we have

$$\frac{\partial f_{b_1}}{\partial s_{ij}} = \sum_{b_2, b_3} \bar{u}_{b_2 b_1} u_{b_3 b_1} (d\mathbf{T})_{b_2 b_3} \left(\frac{\partial}{\partial s_{ij}} \right) = - \sum_{b_2, b_3, b_4, b_5} \bar{u}_{b_2 b_1} u_{b_3 b_1} z_{b_2 b_4} s^{b_4 j} s^{ib_5} \bar{z}_{b_3 b_5},$$

which completes the proof of the last assertion of this lemma. \square

Lemma 10. Let $\Phi(\mathbf{F}) = \text{Diag}(\varphi_1(\mathbf{F}), \varphi_2(\mathbf{F}), \dots, \varphi_m(\mathbf{F}))$, where $\varphi_i(\mathbf{F})$'s ($i = 1, 2, \dots, m$) are differentiable functions from $\mathbb{R}_+^m \rightarrow \mathbb{R}_+$. Then we have

$$\begin{aligned}
\text{Tr} \{ \text{Re} (\nabla'_Z \mathbf{U} \Phi(\mathbf{F}) \mathbf{U}^* \mathbf{Z}) \} &= \sum_k \left\{ f_k \varphi_{kk}(\mathbf{F}) + (p - m + 1) \varphi_k(\mathbf{F}) + 2 \sum_{b > k} \frac{f_k \varphi_k(\mathbf{F}) - f_b \varphi_b(\mathbf{F})}{f_k - f_b} \right\}, \\
\text{Tr} (\mathbf{D}_S \mathbf{Z}^* \mathbf{U} \Phi(\mathbf{F}) \mathbf{U}^* \mathbf{Z}) &= - \sum_k \left\{ f_k^2 \varphi_{kk}(\mathbf{F}) - 2(m - 1) f_k \varphi_k(\mathbf{F}) + 2 \sum_{b > k} \frac{f_k^2 \varphi_k(\mathbf{F}) - f_b^2 \varphi_b(\mathbf{F})}{f_k - f_b} \right\},
\end{aligned}$$

where $\varphi_{kk}(\mathbf{F}) = (\partial \varphi_k / \partial f_k)(\mathbf{F})$, $k = 1, 2, \dots, m$.

Proof. Use notation Φ , φ_k , and φ_{kk} short for $\Phi(\mathbf{F})$, $\varphi_k(\mathbf{F})$, and $\varphi_{kk}(\mathbf{F})$, respectively. To prove the first equation of this lemma, we first note that

$$\text{Tr} \{ \text{Re} ((\nabla'_Z \mathbf{U} \Phi(\mathbf{F}) \mathbf{U}^* \mathbf{Z})) \} = p \text{Tr} (\mathbf{U} \Phi \mathbf{U}^*) + \frac{1}{2} \left(\text{Tr} (\mathbf{Z} \nabla'_Z \mathbf{U} \Phi \mathbf{U}^*) + \text{Tr} (\bar{\mathbf{Z}} \nabla_Z^* \bar{\mathbf{U}} \Phi \mathbf{U}') \right).$$

Now we use the first three equations in Lemma 9 and proceed in a way similar to the proof of

Lemma 8 in order to evaluate the second term inside the expectation of the right hand side as

$$\begin{aligned}
\text{Tr} \{ \mathbf{Z} \nabla'_Z \mathbf{U} \Phi \mathbf{U}^* \} &= \sum_{i,j,k,l} z_{ij} \frac{\partial(u_{kl} \varphi_l \bar{u}_{il})}{\partial z_{kj}} \\
&= \sum_{i,j,k,l} \left\{ z_{ij} \varphi_l \bar{u}_{il} \frac{\partial u_{kl}}{\partial z_{kj}} + z_{ij} \varphi_l u_{kl} \frac{\partial \bar{u}_{il}}{\partial z_{kj}} + z_{ij} u_{kl} \bar{u}_{il} \frac{\partial \varphi_l}{\partial z_{kj}} \right\} \\
&= \sum_{i,j,k,l,b_3,b_5} \left\{ z_{ij} \varphi_l \bar{u}_{il} \sum_{b_1 \neq l} \frac{u_{kb_1} \bar{u}_{kb_1} u_{b_3 l} s^{jb_5} \bar{z}_{b_3 b_5}}{f_l - f_{b_1}} + z_{ij} \varphi_l u_{kl} \sum_{b_1 \neq l} \frac{\bar{u}_{ib_1} \bar{u}_{kl} u_{b_3 b_1} s^{jb_5} \bar{z}_{b_3 b_5}}{f_l - f_{b_1}} \right. \\
&\quad \left. + z_{ij} u_{kl} \bar{u}_{il} \bar{u}_{kb_1} u_{b_3 b_1} s^{jb_5} \bar{z}_{b_3 b_5} \frac{\partial \varphi_l}{\partial f_{b_1}} \right\} \\
&= \sum_{i,b_3,l} \left\{ (\mathbf{ZS}^{-1} \mathbf{Z}^*)_{ib_3} u_{b_3 l} \sum_{b_1 \neq l} \frac{\varphi_l}{f_l - f_{b_1}} \bar{u}_{il} + (\mathbf{ZS}^{-1} \mathbf{Z}^*)_{ib_3} u_{b_3 b_1} \sum_{b_1 \neq l} \frac{\varphi_l}{f_l - f_{b_1}} \bar{u}_{ib_1} \right. \\
&\quad \left. + (\mathbf{ZS}^{-1} \mathbf{Z}^*)_{ib_3} u_{b_3 l} \frac{\partial \varphi_l}{\partial f_l} \bar{u}_{il} \right\} \\
&= \text{Tr} (\mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U} \tilde{\Phi} \mathbf{U}^*),
\end{aligned}$$

where $\tilde{\Phi} = \text{Diag}(\tilde{\phi}_1, \tilde{\phi}_2, \dots, \tilde{\phi}_m)$ with $\tilde{\phi}_i = f_i(\partial \varphi_i / \partial f_i) + \sum_{b \neq i} f_i(\varphi_i - \varphi_b) / (f_i - f_b)$, ($i = 1, 2, \dots, m$). Similarly we use the first three equations in Lemma 9 to evaluate the third term inside the expectation of the right hand side as

$$\begin{aligned}
\text{Tr} \{ \bar{\mathbf{Z}} \nabla_Z^* \bar{\mathbf{U}} \Phi \mathbf{U}' \} &= \sum_{i,j,k,l} \bar{z}_{ij} \frac{\partial(\bar{u}_{kl} \varphi_l u_{il})}{\partial \bar{z}_{kj}} \\
&= \sum_{i,j,k,l} \left\{ \bar{z}_{ij} \varphi_l u_{il} \frac{\partial \bar{u}_{kl}}{\partial \bar{z}_{kj}} + \bar{z}_{ij} \varphi_l \bar{u}_{kl} \frac{\partial u_{il}}{\partial \bar{z}_{kj}} + \bar{z}_{ij} \bar{u}_{kl} u_{il} \frac{\partial \varphi_l}{\partial \bar{z}_{kj}} \right\} \\
&= \sum_{i,j,k,l} \left\{ \bar{z}_{ij} \varphi_l u_{il} \frac{\partial \bar{u}_{kl}}{\partial \bar{z}_{kj}} + \bar{z}_{ij} \varphi_l \bar{u}_{kl} \frac{\partial u_{il}}{\partial \bar{z}_{kj}} + \bar{z}_{ij} \bar{u}_{kl} u_{il} \frac{\partial \varphi_l}{\partial \bar{z}_{kj}} \right\} \\
&= \sum_{i,j,k,l,b_3,b_5} \left\{ \bar{z}_{ij} \varphi_l u_{il} \sum_{b_1 \neq l} \frac{\bar{u}_{kb_1} u_{kb_1} \bar{u}_{b_3 l} \bar{s}^{jb_5} z_{b_3 b_5}}{f_l - f_{b_1}} + \bar{z}_{ij} \varphi_l \bar{u}_{kl} \sum_{b_1 \neq l} \frac{u_{ib_1} u_{kl} \bar{u}_{b_3 b_1} \bar{s}^{jb_5} z_{b_3 b_5}}{f_l - f_{b_1}} \right. \\
&\quad \left. + \bar{z}_{ij} \bar{u}_{kl} u_{il} u_{kb_1} \bar{u}_{b_3 b_1} \bar{s}^{jb_5} z_{b_3 b_5} \frac{\partial \varphi_l}{\partial f_{b_1}} \right\} \\
&= \sum_{i,b_3,l} \left\{ (\mathbf{ZS}^{-1} \mathbf{Z}^*)_{b_3 i} u_{il} \sum_{b_1 \neq l} \frac{\varphi_l}{f_l - f_{b_1}} \bar{u}_{b_3 l} + (\mathbf{ZS}^{-1} \mathbf{Z}^*)_{b_3 i} u_{ib_1} \sum_{b_1 \neq l} \frac{\varphi_l}{f_l - f_{b_1}} \bar{u}_{b_3 b_1} \right. \\
&\quad \left. + (\mathbf{ZS}^{-1} \mathbf{Z}^*)_{b_3 i} u_{il} \frac{\partial \varphi_l}{\partial f_l} \bar{u}_{b_3 l} \right\} \\
&= \text{Tr} (\mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U} \tilde{\Phi} \mathbf{U}^*),
\end{aligned}$$

Putting $\mathbf{ZS}^{-1}\mathbf{Z}^* = \mathbf{U}\mathbf{F}\mathbf{U}^*$ into the right hand side of the above two equations, we have

$$\begin{aligned}
\text{Tr} \{ \text{Re} (\nabla'_Z \mathbf{U} \Phi(\mathbf{F}) \mathbf{U}^* \mathbf{Z}) \} &= \sum_k \left\{ p\varphi_k + f_k \varphi_{kk} + \sum_{b \neq k} \frac{f_k(\varphi_k - \varphi_b)}{f_k - f_b} \right\} \\
&= \sum_k \left\{ p\varphi_k + f_k \varphi_{kk} + \sum_{b \neq k} \frac{f_k \varphi_k - f_b \varphi_b + (f_b - f_k) \varphi_b}{f_k - f_b} \right\} \\
&= \sum_k \left\{ f_k \varphi_{kk} + (p - m + 1) \varphi_k + \sum_{b \neq k} \frac{f_k \varphi_k - f_b \varphi_b}{f_k - f_b} \right\},
\end{aligned}$$

which completes the first equation of this lemma.

To prove the second equation of this lemma, we first note that

$$\begin{aligned}
\text{Tr} (\mathbf{D}_S \mathbf{Z}^* \mathbf{U} \Phi \mathbf{U}^* \mathbf{Z}) &= \sum_{i, j, k_1, k_2, l} \frac{\partial(\bar{z}_{k_1 j} u_{k_1 l} \varphi_l \bar{u}_{k_2 l} z_{k_2 i})}{\partial s_{ij}} \\
&= \sum_{i, j, k_1, k_2, l} \bar{z}_{k_1 j} z_{k_2 i} \left\{ \varphi_l \bar{u}_{k_2 l} \frac{\partial u_{k_1 l}}{\partial s_{ij}} + \varphi_l u_{k_1 l} \frac{\partial \bar{u}_{k_2 l}}{\partial s_{ij}} + u_{k_1 l} \bar{u}_{k_2 l} \frac{\partial \varphi_l}{\partial s_{ij}} \right\}. \quad (27)
\end{aligned}$$

But, from the forth equation of Lemma 9, we have

$$\begin{aligned}
\sum_{i, j, k_1, k_2, l} \bar{z}_{k_1 j} z_{k_2 i} \varphi_l \bar{u}_{k_2 l} \frac{\partial u_{k_1 l}}{\partial s_{ij}} &= - \sum_{b_1 \neq l} \frac{\varphi_l (\mathbf{U}^* \mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U})_{b_1 b_1} (\mathbf{U}^* \mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U})_{ll}}{f_l - f_{b_1}} \\
&= - \sum_{b_1 \neq l} \frac{f_{b_1} f_l \varphi_l}{f_l - f_{b_1}},
\end{aligned}$$

where we denote by $(\mathbf{U}^* \mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U})_{ll}$ the l -th diagonal element of a matrix $\mathbf{U}^* \mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U}$ for $l = 1, 2, \dots, m$. Similarly, from the last two equations of Lemma 9, we have

$$\sum_{i, j, k_1, k_2, l} \bar{z}_{k_1 j} z_{k_2 i} \varphi_l u_{k_1 l} \frac{\partial \bar{u}_{k_2 l}}{\partial s_{ij}} = - \sum_{b_1 \neq l} \frac{f_{b_1} f_l \varphi_l}{f_l - f_{b_1}}$$

and

$$\begin{aligned}
\sum_{i, j, k_1, k_2, l} \bar{z}_{k_1 j} z_{k_2 i} u_{k_1 l} \bar{u}_{k_2 l} \frac{\partial \varphi_l}{\partial s_{ij}} &= \sum_{i, j, k_1, k_2, l, b_1} \bar{z}_{k_1 j} z_{k_2 i} u_{k_1 l} \bar{u}_{k_2 l} \frac{\partial \varphi_l}{\partial f_{b_1}} \frac{\partial f_{b_1}}{\partial s_{ij}} \\
&= - \sum_{b_1, l} (\mathbf{U}^* \mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U})_{b_1 l} (\mathbf{U}^* \mathbf{ZS}^{-1} \mathbf{Z}^* \mathbf{U})_{lb_1} \frac{\partial \varphi_l}{\partial f_{b_1}} \\
&= - \sum_l f_l^2 \frac{\partial \varphi_l}{\partial f_l}.
\end{aligned}$$

Putting these three above equations into (27) we conclude that

$$\begin{aligned}
\text{Tr}(\mathbf{D}_S \mathbf{Z} \mathbf{U} \Phi \mathbf{U}^* \mathbf{Z}^*) &= - \sum_l \left\{ 2 \sum_{b \neq l} \frac{f_b f_l \varphi_l}{f_l - f_b} + f_l^2 \frac{\partial \varphi_l}{\partial f_l} \right\} \\
&= - \sum_l \left\{ 2 \sum_{b \neq l} \frac{\{f_l(f_b - f_l) + f_l^2\} \varphi_l}{f_l - f_b} + f_l^2 \frac{\partial \varphi_l}{\partial f_l} \right\} \\
&= \sum_l \left\{ 2(m-1) f_l \varphi_l - 2 \sum_{b > l} \frac{f_l^2 \varphi_l - f_b^2 \varphi_b}{f_l - f_b} - f_l^2 \frac{\partial \varphi_l}{\partial f_l} \right\},
\end{aligned}$$

which completes the proof of the second assertion of this lemma. \square

FACULTY OF SCIENCE, JAPAN WOMEN'S UNIVERSITY
2-8-1 Mejirodai, Tokyo 112-8681, Japan
email: konno@fc.jwu.ac.jp